

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2005 年 7 月 14 日 (14.07.2005)

PCT

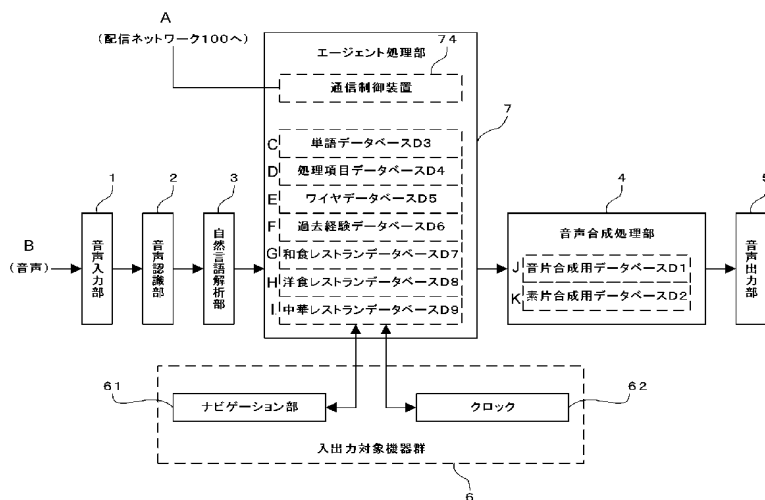
(10) 国際公開番号  
WO 2005/064592 A1

- (51) 国際特許分類<sup>7</sup>: G10L 15/22, (71) 出願人 (米国を除く全ての指定国について): 株式会社ケンウッド (KABUSHIKI KAISHA KENWOOD) [JP/JP]; 〒1928525 東京都八王子市石川町 2 9 6 7-3 Tokyo (JP).
- B60R 16/02, G01C 21/00, G08G 1/0969, G09B 29/10,  
G10L 15/00, 15/18, 15/28
- (21) 国際出願番号: PCT/JP2004/019426 (72) 発明者; および
- (22) 国際出願日: 2004 年 12 月 24 日 (24.12.2004) (75) 発明者/出願人 (米国についてのみ): 佐藤 寧 (SATO, Yasushi) [JP/JP]; 〒2700163 千葉県流山市南流山 4 丁目 1 6 番 1 8 号 4 0 1 号室 Chiba (JP).
- (25) 国際出願の言語: 日本語 (74) 代理人: 木村 満 (KIMURA, Mitsuru); 〒1010054 東京都千代田区神田錦町二丁目 7 番地 協販ビル 2 階 Tokyo (JP).
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願 2003-436976 (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, 2003 年 12 月 26 日 (26.12.2003) JP

[続葉有]

(54) Title: DEVICE CONTROL DEVICE, SPEECH RECOGNITION DEVICE, AGENT DEVICE, ON-VEHICLE DEVICE CONTROL DEVICE, NAVIGATION DEVICE, AUDIO DEVICE, DEVICE CONTROL METHOD, SPEECH RECOGNITION METHOD, AGENT PROCESSING METHOD, ON-VEHICLE DEVICE CONTROL METHOD, NAVIGATION METHOD, AND AUDIO DEVICE CONTROL METHOD, AND PROGRAM

(54) 発明の名称: 機器制御装置、音声認識装置、エージェント装置、車載機器制御装置、ナビゲーション装置、オーディオ装置、機器制御方法、音声認識方法、エージェント処理方法、車載機器制御方法、ナビゲーション方法、オーディオ装置制御方法及びプログラム



A... (TO DISTRIBUTION NETWORK 100)  
B... (SPEECH)  
1... SPEECH INPUT UNIT  
2... SPEECH RECOGNITION UNIT  
3... NATURAL LANGUAGE ANALYSIS UNIT  
7... AGENT PROCESSING UNIT  
74... COMMUNICATION CONTROL DEVICE  
C... WORD DATABASE D3  
D... PROCESS ITEM DATABASE D4  
E... WIRE DATABASE D5  
F... PAST EXPERIENCE DATABASE D6  
G... JAPANESE RESTAURANT DATABASE D7

H... WESTERN RESTAURANT DATABASE D8  
I... CHINESE RESTAURANT DATABASE D9  
4... SPEECH SYNTHESIS PROCESSING UNIT  
J... PHONETIC PIECE SYNTHESIS DATABASE D1  
K... PHONEME PIECE SYNTHESIS DATABASE D2  
5... SPEECH OUTPUT UNIT  
61... NAVIGATION UNIT  
62... CLOCK  
6... I/O OBJECT DEVICE GROUP

(57) Abstract: A speech recognition unit (2) subjects a speech of a speaker inputted to a speech input unit (1) to speech recognition and identifies words and their scores which may be expressed by the speech. A natural language analysis unit (3) identifies the parts of speech of the words and supplies the word data expressing these words to an agent processing unit (7). The agent processing unit (7) contains processing item data defining a data acquisition process for acquiring word data, a judgment process, and an I/O process; and a wire which is data defining transition from a process to another process and giving a weight coefficient to the transition. The processing item data and the wire execute a flow expressed as a whole so as to appropriately grasp the speaker's request and control devices belonging to an I/O object device group (6) so as to satisfy the request.

[続葉有]



BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE,

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約: 音声認識部(2)は、音声入力部(1)が入力した話者の音声に音声認識を施して、この音声が表示している可能性のある単語とそのスコアとを特定し、自然言語解析部(3)がこれらの単語の品詞を特定して、これらの単語を表す単語データをエージェント処理部(7)に供給する。エージェント処理部(7)は、単語データ等を取得するデータ取得処理や、判別処理、入出力処理を定義する処理項目データと、ひとつの処理から次の処理への遷移を定義しこの遷移に重み係数を与えるデータであるワイヤとを記憶しており、処理項目データとワイヤとが全体として表すフローを実行することにより、話者の要求を適切に把握し、この要求を満たすように、入出力対象機器群(6)に属する機器を制御する。

## 明 細 書

機器制御装置、音声認識装置、エージェント装置、車載機器制御装置、ナビゲーション装置、オーディオ装置、機器制御方法、音声認識方法、エージェント処理方法、車載機器制御方法、ナビゲーション方法、オーディオ装置制御方法及びプログラム

### 技術分野

[0001] この発明は、機器制御装置、音声認識装置、エージェント装置、車載機器制御装置、ナビゲーション装置、オーディオ装置、機器制御方法、音声認識方法、エージェント処理方法、車載機器制御方法、ナビゲーション方法、オーディオ装置制御方法及びプログラムに関する。

### 背景技術

[0002] 近年、音声認識の技術を用いて音声を認識し、認識結果に応答して電気機器などを制御する手法が用いられている。この手法は、具体的には、入力した音声が表示単語を識別し、識別された単語が所定のキーワードに合致するか否かを判別して、判別結果に基づいて外部の機器を制御するものである（例えば、特許文献1参照）。  
特許文献1：特開平8-339288号公報

### 発明の開示

#### 発明が解決しようとする課題

[0003] しかし、人間が言語の形で発する指示を完全に認識するのは困難である。このため、上述の手法では、人間が言語の形で発する指示に適切に応答することができない場合があった。

[0004] この発明は上記実状に鑑みてなされたものであり、人間が言語の形で発する指示に適切に応答して機器を制御できる機器制御装置、音声認識装置、エージェント装置、車載機器制御装置、ナビゲーション装置、オーディオ装置、機器制御方法、音声認識方法、エージェント処理方法、車載機器制御方法、ナビゲーション方法、オーディオ装置制御方法あるいはプログラムを提供することを目的とする。

## 課題を解決するための手段

- [0005] 上記目的を達成するため、この発明の第1の観点にかかる機器制御装置は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする。
- [0006] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定してもよい。
- [0007] 前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された複数の語句の組み合わせが所定の条件を満たしているか否かを判別し、判別結果に基づいて、前記音声の発話者の発話内容を特定してもよい。
- [0008] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定してもよい。
- [0009] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定してもよい。
- [0010] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定してもよい。

- [0011] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行してもよい。
- [0012] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促してもよい。
- [0013] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択するようにしてもよい。
- [0014] また、この発明の第2の観点にかかる機器制御装置は、  
 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
 前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
 特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定する処理特定手段(7)と、  
 所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
 前記情報取得手段(7)が取得した情報に基づいて音声を入力する音声入力手段(5)と、  
 を備え、  
 前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を入力するものであるとき、前記音声入力手段(5)は当該情報に基づいて音声を入力する、  
 ことを特徴とする。
- [0015] また、この発明の第3の観点にかかる音声認識装置は、  
 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
 前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内

容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とする。

[0016] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、

前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

[0017] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定するものであってもよい。

[0018] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定するものであってもよい。

[0019] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

[0020] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行するものであってもよい。

[0021] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促すものであってもよい。

[0022] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、

出力する出力内容を選択するものであってもよい。

- [0023] また、この発明の第4の観点にかかる音声認識装置は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、  
ことを特徴とする。

- [0024] また、この発明の第5の観点にかかるエージェント装置は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする。

- [0025] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定

する、  
ことを特徴とする。

- [0026] 前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された複数の語句の組み合わせが所定の条件を満たしているか否かを判別し、判別結果に基づいて、前記音声の発話者の発話内容を特定するものであってもよい。
- [0027] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定するものであってもよい。
- [0028] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定するものであってもよい。
- [0029] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。
- [0030] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行するものであってもよい。
- [0031] 前記特定手段(7)は、入力された音声の意味が判別できなかつたとき、より判別しやすい表現での入力を促すものであってもよい。
- [0032] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択するものであってもよい。
- [0033] 前記処理実行手段(7)は、実行すべき処理として特定した処理が、外部から受信した情報を発話者に提示する処理であるとき、当該情報を読み上げる音声が発生させることにより当該提示を実行する手段を備えるものであってもよい。



[0034] また、この発明の第6の観点にかかるエージェント装置は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、  
ことを特徴とする。

[0035] また、この発明の第7の観点にかかる車載機器制御装置は、  
外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする。

[0036] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品

詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

[0037] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定するものであってもよい。

[0038] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定するものであってもよい。

[0039] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

[0040] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行するものであってもよい。

[0041] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促すものであってもよい。

[0042] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択するものであってもよい。

[0043] また、この発明の第8の観点にかかる車載機器制御装置は、  
外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とする。

[0044] また、この発明の第9の観点にかかるナビゲーション装置は、

車両に搭載可能に構成されたナビゲーション装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定して、当該ナビゲーション処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とする。

[0045] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、

前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

[0046] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定するものであってもよい。

- [0047] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定するものであってもよい。
- [0048] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。
- [0049] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行するものであってもよい。
- [0050] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促すものであってもよい。
- [0051] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択するものであってもよい。
- [0052] また、この発明の第10の観点にかかるナビゲーション装置は、  
車両に搭載可能に構成されたナビゲーション装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、  
を備え、

前記処理特定手段(7)が特定したナビゲーション処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段は当該情報に基づいて音声を出力する、  
ことを特徴とする。

- [0053] また、この発明の第11の観点にかかるオーディオ装置は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、実行すべき音声処理の内容を特定して、当該音声処理を実行し、又は外部の機器に実行させるよう当該外部の機器を制御する処理実行手段(7)と、  
を備える、  
ことを特徴とする。

- [0054] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

- [0055] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定するものであってもよい。

- [0056] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定するものであってもよい。

- [0057] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴ

りに基づいて、前記音声の発話者の発話内容を特定するものであってもよい。

[0058] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行するものであってもよい。

[0059] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促すものであってもよい。

[0060] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択するものであってもよい。

[0061] また、この発明の第12の観点にかかるオーディオ装置は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段と、  
特定した内容に基づいて、実行すべき音声処理の内容を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定した音声処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声出力する、  
ことを特徴とする。

[0062] また、この発明の第13の観点にかかる機器制御方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行ステップと、  
より構成される、  
ことを特徴とする。

[0063] また、この発明の第14の観点にかかる機器制御方法は、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定する処理特定ステップと、

所定の通信装置を介して情報を取得する情報取得ステップと、

前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、  
より構成され、

前記処理特定ステップで特定した制御が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声を出力する、  
ことを特徴とする。

[0064] また、この発明の第15の観点にかかる音声認識方法は、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行ステップと、

より構成される、  
ことを特徴とする。

- [0065] また、この発明の第16の観点にかかる音声認識方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて実行すべき処理を特定する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、  
より構成され、  
前記処理特定ステップで特定した処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声を出力する、  
ことを特徴とする。

- [0066] また、この発明の第17の観点にかかるエージェント処理方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行ステップと、  
より構成される、  
ことを特徴とする。

- [0067] また、この発明の第18の観点にかかるエージェント処理方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、



前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて実行すべき処理を特定する処理特定ステップと、

所定の通信装置を介して情報を取得する情報取得ステップと、

前記情報取得ステップで取得した情報に基づいて音声を入力する音声入力ステップと、

より構成され、

前記処理特定ステップで特定した処理が、前記情報取得ステップで取得された情報を入力するものであるとき、前記音声入力ステップでは当該情報に基づいて音声を入力する、

ことを特徴とする。

[0068] また、この発明の第19の観点にかかる車載機器制御方法は、

車両に搭載された車載機器を制御するための車載機器制御方法であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行ステップと、

より構成される、

ことを特徴とする。

[0069] また、この発明の第20の観点にかかる車載機器制御方法は、

車両に搭載された車載機器を制御するための車載機器制御方法であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定

する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、  
より構成され、  
前記処理特定ステップが特定した制御が、前記情報取得手段で取得された情報を出力するものであるとき、前記音声出力ステップは当該情報に基づいて音声を出力する、  
ことを特徴とする。

[0070] また、この発明の第21の観点にかかるナビゲーション方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定して、当該ナビゲーション処理を実行する処理実行ステップと、  
より構成される、  
ことを特徴とする。

[0071] また、この発明の第22の観点にかかるナビゲーション方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、

より構成され、

前記処理特定ステップで特定したナビゲーション処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声を出力する、  
ことを特徴とする。

[0072] また、この発明の第23の観点にかかるオーディオ装置制御方法は、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、実行すべき音声処理の内容を特定して、当該音声処理を実行し、又は外部のオーディオ装置に実行させるよう当該外部のオーディオ装置を制御する処理実行ステップと、

より構成される、

ことを特徴とする。

[0073] また、この発明の第24の観点にかかるオーディオ装置制御方法は、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、外部のオーディオ装置に実行させるべき音声処理の内容を特定する処理特定ステップと、

所定の通信装置を介して情報を取得する情報取得ステップと、

前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、

より構成され、

前記処理特定ステップで特定した音声処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップで当該情報に基づいて音

声を出力する、  
ことを特徴とする。

- [0074] また、この発明の第25の観点にかかるプログラムは、  
コンピュータを、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする機器制御装置として機能させる。

- [0075] また、この発明の第26の観点にかかるプログラムは、  
コンピュータを、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段と、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声出力する、  
ことを特徴とする機器制御装置として機能させる。

[0076] また、この発明の第27の観点にかかるプログラムは、  
コンピュータを、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段と、  
特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする音声認識装置として機能させる。

[0077] また、この発明の第28の観点にかかるプログラムは、  
コンピュータを、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、  
ことを特徴とする音声認識装置として機能させる。

[0078] また、この発明の第29の観点にかかるプログラムは、  
コンピュータを、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当

該音声が表示する語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とするエージェント装置として機能させる。

[0079] また、この発明の第30の観点にかかるプログラムは、

コンピュータを、

音声を表示する音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表示する語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を入力する音声入力手段(5)と、

を備え、

前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を入力するものであるとき、前記音声入力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とするエージェント装置として機能させる。

[0080] また、この発明の第31の観点にかかるプログラムは、

コンピュータを、

外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、

音声を表示する音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表示する語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、

ことを特徴とする車載機器制御装置として機能させる。

[0081] また、この発明の第32の観点にかかるプログラムは、  
コンピュータを、

外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定する処理特定手段と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、  
を備え、

前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声出力する、

ことを特徴とする車載機器制御装置として機能させる。

[0082] また、この発明の第33の観点にかかるプログラムは、  
コンピュータを、

車両に搭載可能に構成されたナビゲーション装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当

該音声が表示する語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定して、当該ナビゲーション処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とするナビゲーション装置として機能させる。

[0083] また、この発明の第34の観点にかかるプログラムは、

コンピュータを、

車両に搭載可能に構成されたナビゲーション装置であって、

音声を表示する音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表示する語句を特定する音声認識手段と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段と、

特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定する処理特定手段と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を入力する音声入力手段(5)と、

を備え、

前記処理特定手段(7)が特定したナビゲーション処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声入力手段(5)は当該情報に基づいて音声を入力する、

ことを特徴とするナビゲーション装置として機能させる。

[0084] また、この発明の第35の観点にかかるプログラムは、

コンピュータを、

音声を表示する音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表示する語句を特定する音声認識手段(2, 3)と、



前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべき音声処理の内容を特定して、当該音声処理を実行し、又は外部の機器に実行させるよう当該外部の機器を制御する処理実行手段(7)と、

を備える、

ことを特徴とするオーディオ装置として機能させる。

[0085] また、この発明の第36の観点にかかるプログラムは、

コンピュータを、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべき音声処理の内容を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した音声処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とするオーディオ装置として機能させる。

## 発明の効果

[0086] この発明によれば、人間が言語の形で発する指示に適切に応答して機器を制御できる機器制御装置、音声認識装置、エージェント装置、車載機器制御装置、ナビゲーション装置、オーディオ装置、機器制御方法、音声認識方法、エージェント処理方法、車載機器制御方法、ナビゲーション方法、オーディオ装置制御方法あるいはプ

ログラムが実現される。

#### 図面の簡単な説明

- [0087] [図1]この発明の実施の形態に係るエージェント装置を示す図である。
- [図2]この実施の形態に係るエージェント装置のより詳細な構成図である。
- [図3]カテゴリ辞書の具体例を模式的に示す図である。
- [図4]トリガ取得処理を説明するための図である。
- [図5]判別処理を説明するための図である。
- [図6]問い合わせ付きの判別処理を説明するための図である。
- [図7]出力処理を説明するための図である。
- [図8]ワイヤを示す図である。
- [図9]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。
- [図10]重み係数の設定を説明するための図である。
- [図11]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートである。
- [図12]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートの続きである。
- [図13]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートの続きである。
- [図14]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートの続きである。
- [図15]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートの続きである。
- [図16]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートの続きである。
- [図17]搭乗者を食事の場所に誘導するために実行する処理を表すフローチャートの続きである。
- [図18]テーブルのデータ構造を示す図である。

## 符号の説明

- [0088] 1 音声入力部  
2 言語認識部  
3 自然言語解析部  
4 音声合成処理部  
5 音声出力部  
6 入出力対象機器群  
61 ナビゲーション部  
62 クロック  
7 エージェント処理部

## 発明を実施するための最良の形態

[0089] 以下、図面を参照して、この発明の実施の形態を説明する。

### [第1の実施の形態]

[0090] まず、本発明の第1の実施の形態に係るエージェント装置を説明する。このエージェント装置は、入力情報に応じて状態を遷移させて処理を行うものである。また、以下では、このエージェント装置は、車載用のナビゲーション装置として機能するものとする。また、このエージェント装置の一部は、音声認識装置として機能する。

[0091] 図1は、このエージェント装置の構成を示すブロック図である。図示するように、このエージェント装置は、音声入力部1と、音声認識部2と、自然言語解析部3と、音声合成処理部4と、音声出力部5と、入出力対象機器群6と、エージェント処理部7とより構成されている。

[0092] 音声入力部1は、音声を入力し、入力した音声からデジタル形式の音声データを生成し、この音声データを音声認識部2へと供給する。具体的には、音声入力部1は、図2に示すように、例えば、マイクロフォン11と、AF (Audio Frequency) 増幅器12と、サンプルホールド回路を内蔵するA/D (Analog-to-Digital) コンバータ13などより構成されている。マイクロフォン11は音声を音声信号に変換して出力する。AF増幅器12は、マイクロフォン11からの音声信号を増幅して出力する。A/Dコンバータ13は、AF増幅器12からの増幅された音声信号をサンプリング、A/D変換することにより

、デジタル音声データを生成し、音声認識部2へと供給する。

- [0093] 音声認識部2、自然言語解析部3、音声合成処理部4、及びエージェント処理部7は、図2に示すように、それぞれ、例えば、CPU (Central Processing Unit) 等からなるプロセッサ21、31、41、71と、このプロセッサ21、31、41、71が実行するプログラムを記憶するROM (Read Only Memory) やハードディスク装置等の不揮発性メモリ(本実施の形態ではROMを例に説明する) 22、32、42、72と、プロセッサのワークエリアとなる記憶領域を有するRAM (Random Access Memory) 等の揮発性メモリ23、33、43、73とより構成されている。

なお、音声認識部2、自然言語解析部3、音声合成処理部4、及びエージェント処理部7の一部又は全部の機能を1つのプロセッサや1つの不揮発性メモリや1つの揮発性メモリで構成してもよい。

- [0094] 音声認識部2は、音声入力部1より供給された音声データに音声認識処理を行う。

音声認識部2は、音声入力部1より音声データを供給されると、例えばこの音声データに含まれる雑音を除去する処理等を行った後、後述するエージェント処理部7内の単語データベースD3を参照して、この音声データに音声認識を施す処理を行うことにより、この音声データが表している可能性のあるすべての単語の候補と、この候補の尤度(スコア)Sと、を特定する。

- [0095] 音声認識部2は、具体的には、例えば、入力された音声と単語データベースD3に登録されるすべての単語とのマッチングの度合い、即ち尤度(スコア)Sを算出して、どの単語のスコアSが最も高いかを認識する。例えば、「えーと、???が減ったよ」という文章を表しており、ただし「???」の部分が「お腹」とも「お墓」とも認識され得るような音声を表す音声データが供給された場合は、「???」の部分については、単語「お腹」に対するスコアSが80%、単語「お墓」に対するスコアSが65%、などと、それぞれの単語に対するスコアSを計算する。そして、最も高い値を示した「お腹」を、当該部分「???」が表す単語であるとして選択する。そして、音声認識部2は、「えーと、お腹が減ったよ」という認識結果を示すテキスト文字列(若しくは単語ID (Identifier))を生成して、自然言語解析部3に供給する。

- [0096] なお、音声認識部2が行う音声認識の手法は任意である。また、スコアSの値が所

定値を下回る単語については、候補として選択しないようにしてもよい。さらに、1つの音声入力に対して単語の候補が複数特定されてもよい。そして、特定した候補及び当該候補のスコアSを示すデータ(以下、単語データと呼ぶ)を生成し、エージェント処理部7へと供給する。尚、単語データは、単語情報そのものとスコアSを含むものとしても良いが、実際には単語IDとスコアSを含むものとした方が、データの取扱上有利である。

[0097] 自然言語解析部3は、音声認識部2より供給された単語データ(若しくはテキストデータ)に形態素解析を施す処理を行うなどすることにより、この単語データが表している単語を品詞ごとに分類し、分類結果を示すデータを単語データに付してエージェント処理部7へと供給する。なお、自然言語解析部3が行う分類の手法は任意であり、例えば、奈良先端科学技術大学で開発された日本語形態素解析の手法である“ChaSen”などを用いればよい。

[0098] 音声合成処理部4の不揮発性メモリ42は、単語の波形を表すデータを記憶する音片データベースD1と、音素を構成するための波形を表すデータを記憶する素片データベースD2とを記憶する。

音片データベースD1は、単語の波形を表すデータを記憶する。素片データベースD2は、音素を構成するための波形を表すデータを記憶する。音声合成処理部4は、音片データベースD1及び／又は素片データベースD2に格納されているデータを用いて、エージェント処理部6より供給された文章データを読み上げる音声を表すデジタル音声データを生成する。

音声合成処理部4は、生成した音声データを音声出力部5に供給する。

デジタル音声データを生成する手法は任意であるが、例えば、録音編集方式や規則合成方式(Rule-based synthesis)を使用できる。なお、録音編集方式は、例えば、単語単位若しくは文節単位の音声を予めアナウンサーに読んでもらい、それらをつなぎあわせて出力する方式である。若しくは、アナウンサーが読み上げた一連の音声を、後の処理で単語単位若しくは文節単位に区切って記憶し、それらをつなぎあわせて出力するようにしてもよい。また、規則合成方式は、音韻(子音や母音)や仮名、更には、より細分化された単位の音素や素片のような比較的小さな単位をつなぎあ

わせて出力する方式である。

- [0099] 音声出力部5は、音声合成処理部4から供給されたデジタル音声データが表す音声再生する。より詳細には、音声出力部5は、図2に示すように、D/A(Digital-to-Analog)コンバータ51、AF増幅器52及びスピーカ53を備える。

D/Aコンバータ51は、音声合成処理部4より供給されたデジタル音声データをD/A変換して、アナログ音声信号に変換する。AF増幅器52は、アナログ音声信号を増幅する。スピーカ53は、アナログ音声信号に従って振動し、アナログ音声データが表す音声を再生し、放音する。

- [0100] 入出力対象機器群6は、例えば、ナビゲーション部61や、クロック62などより構成されている。

- [0101] ナビゲーション部61は、GPS(Global Positioning System)の移動局や液晶ディスプレイ等の表示装置やプロセッサを用いた公知のカーナビゲーションシステムなどより構成されている。ナビゲーション部61は地図を表す地図情報を記憶しており、エージェント処理部7が供給する制御信号に従って車両の現在位置を検知し、検知した現在位置が地図上でどの位置に当たるかを表すデータを生成して、現在位置近隣の地図を表す地図情報と共にエージェント処理部7に供給する。また、現在位置近隣の地図を、現在位置を示す図形と共に表示する。

- [0102] クロック62は、例えば水晶発振器やカウンタ回路等より構成されており、現在の時刻を示すデータを連続的に生成し、エージェント処理部7へと供給する。なお、ナビゲーション部61を介してGPSから得られる時刻情報を供給するようにしてもよい。

- [0103] エージェント処理部7は、モデムやパケット通信端末等からなる通信制御装置74を備え、この通信制御装置を介して外部のネットワーク(例えば、無線電話回線を介したインターネット)を介して後述する配信サーバ100に接続している。

- [0104] エージェント処理部7の不揮発性メモリ72は、単語データベースD3を記憶する。単語データベースD3は、複数の単語のデータと、この単語がどのような概念ないしカテゴリの下にグルーピングされているかを示すためのカテゴリ辞書とを格納するデータベースである。

- [0105] カテゴリ辞書は、例えば図3に示すデータ構造を有する。すなわち、カテゴリ辞書は

、単語を示すデータと、当該単語が属するカテゴリを示すデータとを、互いに対応付けて格納している。例えば、図3に示す例では、単語「減った」はカテゴリ「食事」に対応付けられている。なお、1個の単語が複数のカテゴリに属していてもよい。(例えば図3に示す例では、単語「お腹」は、カテゴリ「食事」及び「病院」に対応付けられている。)

[0106] そして、このエージェント装置は、複数の単語が同一の文章に含まれていた場合に、これらの単語が共通して属するカテゴリのみを、これらの単語が属するカテゴリとして扱うことにより、これらの単語が発話された意図を文脈に応じて把握することを可能とする。例えば図3に示す例では、「お腹が減った」という文章を構成する単語データの集合が自然言語解析部3からエージェント処理部7に供給された場合、エージェント処理部7が、単語「お腹」及び「減った」が共通して属するカテゴリ、すなわち「食事」のみを、当該文章に係る処理を行う限りでは、単語「お腹」及び「減った」が属するカテゴリとして扱うこととできる。同様に、単語データの集合が、ユーザの発話「食事に行きたいなあ」を表している場合は、単語「食事」及び「行きたい」が共通して属するカテゴリ「食事」のみを、文章「食事に行きたいなあ」に係る処理を行う限りでは、単語「食事」及び「行きたい」が属するカテゴリとして扱うこととできる。

[0107] また、エージェント処理部7は、自然言語解析部3から供給された単語データのうち、特定の品詞を表すもののみ(例えば、名詞と動詞のみ)を、自己が行う後続の処理に用いるようにしてもよい。

[0108] なお、単語が属するカテゴリは必ずしも複数の単語の組み合わせから特定される必要はない。例えば、単語「すいた」はカテゴリ「食事」にのみ対応付けられているから、単語「すいた」を表す単語データが供給された場合、エージェント処理部7は、他の単語と組み合わせで判別処理を行うまでもなく、この単語データをカテゴリ「食事」に属するものとして扱ってよい。

[0109] また、エージェント処理部7の不揮発性メモリは、更に、過去経験データベースD6、和食レストランデータベースD7、洋食レストランデータベースD8及び中華レストランデータベースD9を記憶している。過去経験データベースD6は、エージェント処理部7が、ナビゲーション部61に地図上の位置を表示させたり、目的地として設定された

レストランの名称を格納するデータベースである。和食レストランデータベースD7、洋食レストランデータベースD8及び中華レストランデータベースD9は、それぞれ、和食レストランの名称、洋食レストランの名称、及び中華レストランの名称を格納するデータベースである。

[0110] エージェント処理部7の不揮発性メモリは、更に、処理項目データベースD4及びワイヤデータベースD5を記憶している。

[0111] 処理項目データベースD4は、エージェント処理部7が行うトリガ取得処理(TGxx)、判別処理(BRxx又はQBxx)及び入出力処理(後述するSPxx又はEXxx)の内容を、処理項目(ポイント)毎に記述するデータ(処理項目データ)を格納したデータベースである。なお、「xx」は識別番号である。

[0112] 処理項目データベースD4に格納される処理項目のうち、「トリガ取得処理(TGxx)」の内容を記述するデータは、これらの処理を開始させるトリガを特定するトリガデータ(トリガとして取得するデータの内容を指定するデータ)と、後述する進行方向決定用の遷移定数(その進行方向に遷移する程度を示すもので、後述する重み係数Jの算出基準となる定数)とを含む。トリガデータは任意であり、例えば、ナビゲーション部61より供給される、車両の現在位置を示すデータや、クロック62より供給される、現在時刻を示すデータや、自然言語解析部3より供給される上述の単語データである。あるいは、トリガデータは、エージェント処理部7自身が行う処理から引き渡されるデータであってもよい。また、トリガ取得処理で取得されるデータが単語データである場合は、当該単語データが表す単語に代えて、当該単語データが表す単語がグルーピングされているカテゴリが記述されてもよい。ただし、トリガ取得処理の内容は、複数のトリガ取得処理が互いに同一の単語を表す単語データや互いに同一のカテゴリに属す単語データに基づいて動作することがないように記述されるものとする。このように記述しないと、ある単語データによって複数の動作を実行しようとしてしまうことになる。

図4(a)にトリガ取得処理TGxxの例を示す。この例では、トリガTG01は、トリガとしてのカテゴリ「食事」を取得する(カテゴリ「食事」にグルーピングされている単語(図3の例では、単語「減った」、「すいた」、「お腹」、「食べ」、「行く」、「設定」、「どこか」、「



どこに」、「食事」)を識別する)処理であり、その処理に後続する処理に進む(遷移する)か否かを決定するための遷移定数 $k$ は0.8である。図4(b)には、トリガ取得処理TG01のフローチャートを示す。

トリガ取得処理TG00は単語「つまない」を取得する処理である。トリガ取得処理TG51は単語「和食」を取得する処理である。トリガ取得処理TG52は単語「洋食」を取得する処理である。トリガ取得処理TG53は単語「中華」を取得する処理である。

[0113] 処理項目データベースD4に格納されている「判別処理(BRxx)」の内容を記述するデータは、判別条件と、判別結果としてとり得る結果のリストと、後述する戻り方向の遷移定数 $k$ とを、判別処理別に記述したデータを含んでいる。判別処理の内容を記述するデータは、進行方向を決定するための遷移定数 $k$ を、判別結果毎に記述したデータを含んでいる。

図5(a)に判別処理BRxxの例を示す。この例では、判別処理BR01は、「(a)12時より前、(b)12時以降で14時より前、(c)14時以降、のいずれであるかを判別する」である。(a)12時より前と判別したときに後続する処理に進むか否かを決定するための遷移定数 $k$ が0.4、(b)12時以降で14時より前と判別したときに後続する処理に進むか否かを決定するための遷移定数 $k$ が0.3、(c)14時以降と判別したときにそれに後続する処理に進むか否かを決定するため遷移定数 $k$ が0.4である。この例のフローチャートを図5(b)に示す。図5(b)に示すノードBR01.1は、処理の開始点を示す始点ノード、ノードBR01.2は、(a)12時より前と判別したときの進行方向のノードであり、その遷移定数 $k$ は0.4である。さらに、ノードBR01.3は、(b)12時以降で14時より前と判別したときの進行方向のノードであり、その遷移定数 $k$ は0.3であり、ノードBR01.4は、(c)14時以降と判別したときの進行方向のノードであり、その遷移定数 $k$ は0.4である。

[0114] 「判別処理」は、判別に用いるデータを任意の取得源から取得する場合があってもよいとする。取得源としては、例えば、音声認識部2や、自然言語解析部3、エージェント処理部7が実行する他の処理や、入出力対象機器群6に属する機器や、その他外部の機器などが考えられる。そしてこの場合、判別処理の内容を記述するデータは、例えば、判別に用いるデータの取得源を指定するデータを更に含んでいればよ

い。

- [0115] また、「判別処理」では、所定のデータを、判別に先立って所定の出力先に出力するようにしてもよい(この場合は、処理を示す記号を例えばQBxxとする)。例えば、所定の質問を表すデータを、判別に先立って音声合成処理部4に引き渡す、等が考えられる。判別処理において所定のデータを判別に先立って所定のデータを出力する場合、判別処理の内容を記述するデータは、例えば、出力するデータの内容と、このデータの出力先とを指定するデータを含む。

図6(a)に判別処理QBxxの例を示す。この例では、例えば、判別処理QB01は、「食事に行く?」と利用者に問い合わせ、その応答(利用者の回答)が「はい」であったときの進行方向の遷移定数kが0.7、「いいえ」であったときの進行方向の遷移定数kが0.4である。この例のフローチャートを図6(b)に示す。図6(b)に示すノードQB01.1は、処理の開始点を示す始点ノード、ノードQB01.2は、問い合わせに対して、「食事に行く」ことが指定されたことを判別したときの進行方向のノードであり、その遷移定数kは0.7である。さらに、ノードQB01.3は、「食事に行かない」ことが指定されたことを判別したときの進行方向のノードであり、その遷移定数kは0.4である。また、判別処理QB02は、「それじゃコンビニに行く?」と利用者に問い合わせ、その応答(利用者の回答)が「はい」であったときの進行方向の遷移定数kが0.5、「いいえ」であったときの進行方向の遷移定数kが0.3である。

- [0116] 処理項目データベースD4に格納されている「入出力処理」の内容を記述するデータは、入力あるいは出力するデータの内容を指定するデータから構成されている。入力データ及び出力データは任意の内容を有してよい。例えば、出力データは、音声合成処理部4を介して音声出力部5に発生させる音声の読みを表すデータや外部の機器を制御する制御信号であってもよい。また、入力データは、例えば、外部の機器から供給されるデータであってもよい。

図7(a)に出力処理EXxxの例を示す。この例では、例えば、出力処理EX02は、「ナビゲーション案内を行う」という動作であり、動作後の処理を行う進行方向の遷移定数kが0.8である。この例のフローチャートを図7(b)に示す。図7(b)に示すノードEX01.1は、処理の開始点を示す始点ノード、ノードEX01.2は、処理の終了を示す

ノードであり、遷移定数 $k$ は0.8である。尚、出力処理EX<sub>xx</sub>については、遷移定数 $k$ の設定等を行わずに、処理の終了を示すノードの選択を必須の処理としてもよい。

[0117] ワイヤデータベースD5は、複数の処理(TG、BR<sub>xx</sub>、QB<sub>xx</sub>、SP<sub>xx</sub>、EX<sub>xx</sub>)間の遷移を記述するデータ(以下、この遷移定義データをワイヤと呼ぶ)の集合から構成されている。ワイヤは、例えば図8に示すような書式で記述されたデータから構成されている。ワイヤW<sub>n</sub>(W1、W2...)は図示するように、先行する処理X(From(X))から後続する処理Y(To(Y))への遷移(From(X) To(Y))について、当該先行の処理(X)と、当該後続の処理(Y)と、当該遷移に対して与えられた重み係数 $J$ と、を指定するデータである。なお、先行の処理Xが判別処理である場合は、当該判別処理のどの判別結果からの遷移であるか、まで記述される必要がある。具体的には、例えば、質問への回答を音声データに基づいて判別するような処理項目を先行の処理項目として有する1つのワイヤに対して、肯定を意味する単語「行く」「うん」「そうだね」が定義され、また他のワイヤに対しては、否定を意味する単語「行かない」「いや」等が定義される、等である。

[0118] なお、上述した処理項目データが記述する遷移定数 $k$ が、接続されるワイヤが示す遷移が実行された実績に従ってエージェント処理部7等により書き換えられるようにしてもよい。例えば、特定のワイヤが示す遷移が実行される頻度が所定量より大きい場合、エージェント処理部7が、このワイヤが示す遷移の遷移元の処理項目を示す処理項目データに記述されている遷移定数 $k$ の値を従前より大きな値へと書き換えることによって、このワイヤが示す遷移が起りやすくなる、等である。こうすることにより、結果的にユーザの意図する応答がなされる確率が高くなる。

[0119] そして、エージェント処理部7は、処理項目データベースD4及びワイヤデータベースD5が全体として表しているフローを実行する。処理項目データベースD4及びワイヤデータベースD5は、例えば図9(a)に示すようなフローを記述することができる。

[0120] ここで、エージェント処理部7の処理が、具体的にどのように行われるかについて、図9(a)を参照して詳述する。エージェント処理部7は、図示するように、先行する第1の処理P1を実行して後続の第2の処理P2に遷移するようにワイヤW01により定義されており、また、第2の処理P2を実行して後続の第3の処理P3に遷移するようにワイ

ヤW03によって定義されているとき、以下の処理を行う。

[0121] なお、図示するように、処理P1は、単語「行く」を表す単語データが供給されたか否かを判別する処理であるとし、処理P2は、単語「食事」を表す単語データが供給されたか否かを判別する処理であるとし、処理P3は、レストラン又は病院のどちらの位置を示す情報をナビゲーション部61から取得するか否かを判別する処理であるとする。なお、図示するように、処理P1〜P3のいずれについても、各進行方向の遷移定数 $k$ はいずれも0.5であるものとする。この場合、ワイヤは、例えば、図9(b)に示すように、定義される。また、エージェント処理部7は、音声認識部2が自然言語解析部3に供給した単語データのうち、自然言語解析部3において、名詞又は動詞であると分類された単語を表す単語データのみを取得するものとする。

[0122] まず、エージェント処理部7が第1の処理P1に到達しているとき、エージェント処理部7は、ワイヤW01、W03及びW05のそれぞれの重み係数 $J$ を計算し、計算結果をワイヤW01、W03及びW05に書き込む。これら値は、各処理に予め設定されている進行方向の遷移定数 $k$ によって決定される。

[0123] 具体的には、処理P1に処理が到達したとき、ワイヤW01の重み係数 $J$ は、処理P1のワイヤに係る遷移定数 $k$ の値すなわち0.5となる。

処理P2のワイヤW03の重み係数 $J$ は、処理P1のワイヤW01に係る遷移定数 $k=0.5$ に処理P2のワイヤW03に係る遷移定数 $k=0.5$ を乗じた結果すなわち0.25となる。

ワイヤW05の重み係数 $J$ は、処理P1のワイヤW01に係る遷移定数 $k=0.5$ に処理P2のワイヤW03に係る遷移定数 $k=0.5$ を乗じた結果に更に処理P3のワイヤW05に係る遷移定数 $k=0.5$ を乗じた結果、すなわち0.125となる。

[0124] このようにして、ある処理を基点としたときのそれぞれのワイヤの重み係数 $J$ が計算される。よって、現在の状態が遷移すると、現在の処理を基点にしてその都度重み係数 $J$ が計算されることになる。

[0125] 具体的には、現在の状態が処理P2に遷移すると、ワイヤW03の重み係数 $J$ は、処理P2のワイヤW03に係る遷移定数 $k$ に等しい値0.5となり、ワイヤW05の重み係数 $J$ は処理P2のワイヤW03に係る遷移定数 $k=0.5$ と処理P3のワイヤW05に係る遷移

定数 $k=0.5$ との積すなわち $0.25$ となる。またこのとき、エージェント処理部7は、逆方向、つまり処理P1に戻る方向に係るワイヤW01の重み係数Jも再度書き込まれる。処理P2に遷移した場合は、ワイヤW01に係る戻り方向の遷移定数 $k=0.1$ がそのままワイヤW01の重み係数Jとなる。処理P3に遷移した場合は更に、ワイヤW03に係る戻り方向の遷移定数 $k=0.1$ がそのままワイヤW03の重み係数Jとなる。そして、処理P3に遷移した状態におけるワイヤW01の重み係数Jは、処理P3に遷移した状態におけるワイヤW03の遷移定数 $k=0.1$ に、処理P2の戻り方向の遷移定数 $k=0.1$ を乗じた値すなわち $0.01$ となる。

各ワイヤWnの重み係数Jの変化を図9(c)に示す。

[0126] 重み係数Jの計算は、関連するフローの処理のみではなく、すべてのフローのすべてのワイヤについて設定される。ここで現在の処理に関連のないワイヤについては、予め定められた低い計数値を割り当てるようにすればよい。しかし、特にトリガ取得処理を先行の処理とするワイヤについては、重み係数Jをある程度高く設定するようにする。こうすることによって、直前までなされていた会話と著しく異なる内容の会話にもジャンプすることが可能になる。

[0127] なお、各ワイヤには条件がそれぞれ設定されている。具体的には、W01には単語「行く」が条件として設定され、W03には単語「食事」が条件として設定される。そして、自然言語解析部3から単語データが供給されたとき、その単語データが単語「行く」を示すものである場合には、当該単語データが示すスコアSをW01に対して設定する。また、自然言語解析部3から供給された単語データが単語「食事」を示すものである場合には、当該単語データが示すスコアSをW03に対して設定する。尚、各ワイヤに対して異なる単語を複数、条件として設定してもよい。例えば、ある一つのワイヤに対して、「食事」を意味する単語「食事」「めし」等を条件として割り当てると共に、「行く」を意味する単語「行く」「する」等を条件として割り当てる。そして、自然言語解析部3から供給された単語データが、単語「食事」と単語「行く」であった場合、それぞれのスコアSが当該ワイヤに設定される。そして、この設定されたスコアSに基づいて当該ワイヤの算出結果を求める。この場合、各スコアSを合算してもよいし、または平均値を求めてもよい。

[0128] 更に、各ワイヤに設定される条件は、ある意味に対して単一の単語が設定される場合に限られない。例えば、同一の意味を表す異なる複数の単語が条件として設定されるようにしてもよい。この条件の設定は、これらの単語とワイヤの関係が記憶されていれば足り、スコアS算出の対象となる単語は前述の単語データベースD3に記憶される。

[0129] 例えば、音声認識を行う場合に、入力された音声について、音声認識部2は単語データベースD3に登録されているすべての単語に対してスコアSの計算を行い、少なくとも所定以上のスコアが得られた単語に関する単語データを生成して、自然言語解析部3を介してエージェント処理部7に単語データを出力する。次に、エージェント処理部7は、入力された各単語データが、いずれのワイヤに関連付けられているかを判別し、単語データが示すスコアSを関連する各ワイヤに設定する。このようにすれば、複数のワイヤに対して、同じ単語が条件として設定されている場合でも、入力された音声信号と単語とのスコアSの計算は一度で足りることになる。そして、得られた各単語のスコアSが、それぞれ関連付けられているワイヤSに対して設定される。なお、1つのワイヤSに対して複数のスコアSが得られる場合が生じることになるが、この場合には、例えば最も値が高いスコアSを選択するようにすればよい。

[0130] これによって、例えば、「イエス」「はい」「そうだね」等、同一の意味を示す単語データがそれぞれ1つのワイヤに対して設定されていれば、ユーザが「イエス」「はい」「そうだね」のいずれの音声で指示を出しても、音声認識のスコアSとして適切な結果を得ることができる。

[0131] また、判別処理によっては、入出力対象機器群6からの情報が入力情報として得られる場合がある。この場合、例えば音声認識部2を構成するプロセッサ21が、上述の不揮発性メモリ等に、入出力対象機器群6に係る各機器の状態を示すデータを、(例えば上述の「単語データベースD3」と同様の)状態データベースとして格納し、入力情報がどの機器のどの状態を示すものかを状態データベースを参照して決定し、これとスコアSとを状態データとしてエージェント処理部7に供給すればよい。なお、音声認識の場合と異なり、対応する状態が状態データベースにあれば、スコアSは必ず100%になるものとすればよい。そして、エージェント処理部7は、状態データが示す

状態に関連するワイヤを判別し、各ワイヤに対してスコアSを設定すればよい。

[0132] 次に、このように構成されたシステム全体の動作を、図10を参照して説明する。

音声入力部1と音声認識部2と自然言語解析部3とは、独自に動作して、音声を取り込み、解析し、単語データをエージェント処理部7に提供する。

そして、エージェント処理部7は、判別条件に係る(1個又は複数個の)単語データ(若しくは状態データ)が自然言語解析部3等より供給されると、以下の処理を行う。

供給された単語を認識(識別)し(ステップS11)、それが、単語データベースDB4に登録されている単語に相当するか否かを判別する(ステップS12)。登録されていなければ(ステップS12, No)、単語入力処理を終了する。

一方、登録されていれば(ステップS12, Yes)、その単語又はその単語の属す「カテゴリ」が条件となっている処理に関し、単語の尤度つまりスコアSとワイヤの重み係数Jの積 $S \cdot J$ を計算する(ステップS13)。

例えば、図9(a)に示すフローを実行している場合において、処理ポインタPPが第1の処理を指示しているとする。この場合の、各ワイヤの重み係数Jは図9(c)に示すとおりである。

この状態で、スコアSが80%の単語「行く」と、スコアSが50%の単語「食事」を示す単語データが入力されたと仮定する。

図9(a)と(b)に示す例では、第1の処理P1では、単語「行く」が判別条件に関連し、第3の処理P3では、単語「食事」が判別条件に関連する。

図9(c-1)示すように、単語「行く」を示す単語データを入力する処理を先行の処理とするワイヤW01の重み係数Jが0.5、単語「食事」を示す単語データを入力する処理を先行の処理とするワイヤW03の重み係数Jが0.25である。この場合、ワイヤW51及びW53について求められる尤度Sと重み係数Jの積 $S \cdot J$ は、数式1及び2に示すとおりとなる。

[0133] (数1) ワイヤW01についての判別結果:「行く」に対するスコア $S (=80\%) \times$ ワイヤW01の重み係数 $J (=0.5) = 40$

[0134] (数2) ワイヤW03についての判別結果:「食事」に対するスコア $S (=50\%) \times$ ワイヤW03の重み係数 $J (=0.25) = 12.5$

[0135] エージェント処理部7は、スコアSと重み係数Jとの積 $S \cdot J$ を求める上述の処理を、フローが有するすべてのワイヤについて行う。

続いて、エージェント処理部7は、計算された積 $S \cdot J$ がもっとも大きいワイヤを選択する(図10のステップS14)。エージェント処理部7は、選択したワイヤに後続する処理に制御を進める(ステップS15)。例えばワイヤW01について求めた積 $S \cdot J$ が最も高い値を示した場合、入力された単語データは単語「行く」を示すものであったと認識して、ワイヤW01が後続の処理としている第2の処理P2に遷移する。通常、現在の実行中の処理を起点するとワイヤの重み係数Jが比較的大きい。このため、一般的には、次の処理に移るが、従前と全く異なる単語で尤度Sの高いものが入力された場合には、その単語に対応する処理が開始されることもある。例えば、上述の例でいえば、「行く」に対するスコアSが30%であり、「食事」に対するスコアSが80%であった場合、それぞれの積 $S \cdot J$ は15と20になり、この場合には、ワイヤW03が選択されることになる。これは、当該エージェント装置のユーザが処理の流れを熟知しており、現在の処理を飛ばして近接する他の処理にジャンプさせる場合などに特に有効である。

[0136] エージェント処理部7は、処理が遷移すると、遷移後の状態に基づいて、各ワイヤの重み係数Jを再計算する(ステップS16)。

以後は、その処理の内容に従って処理を進める(ステップS17)。この例では、第2の処理P2を実行する。

[0137] なお、トリガ取得処理からの遷移に対しては、ある程度高い重み係数Jを設定しておくといふ。具体的には、例えば、処理P1には「行く」という単語を示す単語データを取得するトリガ取得処理が先行しているものとして、このトリガ取得処理から処理P1への遷移を定義するワイヤに対しては、例えば重み係数 $J = 1.0$ を与えておく。そうすると、例えばエージェント処理部7の処理がこのトリガ取得処理に係るフローに係属している場合において、ユーザが「食事に行く」と発音し、例えば単語「行く」に対するスコアSが80%である単語データが得られれば、このスコアSと、「ユーザが「行く」と言ったか否かの判断」に係るこのワイヤの重み係数Jとの積 $S \cdot J$ は、 $80\% \times 1.0$ すなわち80となる。この値が他のワイヤの判別結果と比べて大きな値であれば、入力された音声は「行く」であったと認識され、エージェント処理部7の処理が処理P1にジャンプされ



る。一方で、他のワイヤの重み係数 $J$ を極めて低く設定しておけば、これら他のワイヤにより定義されている処理にジャンプされる可能性は極めて低くなり、結果として、ある程度想定される会話の流れに沿って認識率を向上させることができる。

- [0138] この実施の形態では、戻り方向への遷移も起こり得る。しかし、現実的には会話を戻すことは好ましくないことが多い。そこで、戻り方向の遷移定数 $k$ は、進行方向の遷移定数 $k$ に比べて低い値に設定するようにすればよい。そうすると、入力された音声から高いスコア $S$ の単語データが仮に得られても、戻り方向の遷移定数 $k$ が重み係数 $J$ として書き込まれたワイヤについて求めた積 $S \cdot J$ は低い値となるため、戻り方向への遷移の可能性を低く抑えることができる。

また、エージェント処理部7は、求めた積 $S \cdot J$ の値が所定の条件に合致しないような処理(たとえば、積 $S \cdot J$ の値が所定値に達しないような処理)は、遷移を実行する対象から除外するよう取り扱ってもよい。

- [0139] なお、例えば図8に示しているように、ワイヤは、処理項目から処理項目への遷移という形で遷移を定義する。そして、ワイヤを図8に示すような形態で記述してデータベースに格納することにより、各処理項目同士の関係を、あたかもコンピュータのマクロ処理のように定義することが可能になる。これによって、各処理項目を容易に接続することができる。

- [0140] また、トリガとなる処理項目は、実際には接続されるワイヤに係る条件としての単語等(他の入力対象機器群6に対する状態の場合もあり得る)のスコア $S$ と重み係数 $J$ に基づく判別結果の算出になるので、ワイヤの定義においてトリガ取得処理項目はワイヤの開始点として定義されず、ワイヤそのものが遷移元として定義されることになる。

- [0141] 更に、上述のように各処理項目の接続関係をワイヤによって定義しているので、簡単に遷移先を追加することができる。例えば、「暑い」という音声入力の後に、ユーザが休憩することを意図して「ファミリーレストランを探して」という音声を入力する機会が多い場合、ファミリーレストランの検索処理項目に対して自動でワイヤを追加する。そうすると、ワイヤが自動で追加された後には、ファミリーレストラン検索処理項目に接続されたワイヤの重み係数 $J$ をある程度大きくすることで、当該入力「ファミリーレストランを探して」に適切に対応することができるようになる。(ただしこの場合、エージェント

処理部7は、例えばファミリーレストランの位置を示す情報を含んだ地図データ等を記憶し、あるいは外部の地図データ等にアクセスするものとする。)

このワイヤの自動追加は、ある処理項目からある処理項目(若しくはワイヤ)へのジャンプの回数を計数し、これが所定回数に達したときに自動で行うようにすればよい。

[0142] 次に、このエージェント装置が、ユーザである搭乗者を食事の場所に誘導するために実行する処理を、図11〜図17を参照して説明する。以下では、処理項目データベース及びワイヤデータベースは、例えば、全体として図11〜図17に示すようなフローを記述しているものとする。

[0143] 図11〜図17に示すフローにおいては、エージェント処理部7は、トリガ取得処理ステップTG1ではカテゴリ「食事」に属する単語を示す単語データが自然言語解析部3より供給されると、これを取得して判別処理ステップQB1に引き渡す。トリガ取得処理ステップTG1で取得される単語は「食事」のカテゴリに分類されている単語であればよいので、例えば「ごはん」、「食べ物」など任意の単語をカテゴリ辞書によって「食事」のカテゴリに分類しておけば、このエージェント装置は、様々な単語に応答して判別処理ステップQB1以下の処理を実行するようになる。

[0144] エージェント処理部7は、判別処理ステップQB1では、まず音声合成処理部4に、「食事に行く?」という文章を表す単語データを供給する。音声合成処理部4は、この文章を読み上げる音声を表す音声データを生成して音声出力部5に供給し、音声出力部5は、この音声データが表す音声を再生する。そしてエージェント処理部7は、自然言語解析部3から、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、該当する単語データが供給されると、このデータが「肯定」又は「否定」のいずれに属するかを判別する。そして、「肯定」に属すると判別すると判別処理ステップQB3に処理を移し、「否定」に属すると判別すると判別処理ステップQB2に処理を移す。

尚、カテゴリに基づく処理ではなく、各ワイヤ、例えばQB1からQB3への状態遷移を定義するワイヤに対しては、肯定を意味する単語「はい」「そうだね」等を条件として設定し、また、QB1からQB2への状態遷移を定義するワイヤに対しては、否定を意味する単語「いや」「行かない」等を条件として設定するようにしてもよい。これによっ

て、自然言語解析部3から入力音声に対応する単語のスコアSが直接ワイヤに対して設定される。例えば、入力音声の単語「そうだね」に対するスコアSが80%であれば、このスコアSがQB1からQB3への状態遷移を定義するワイヤに対して設定されることになる。

そして、上述のカテゴリに各単語を分類する場合でも同様だが、ここで得られた各スコアSとワイヤに設定される重み係数Jとの乗算によってワイヤの算出結果が求められ、この算出結果が最も高い値を示すワイヤに状態を遷移することになる。

また、「肯定」又は「否定」のいずれのカテゴリに属する単語の単語データも供給されなかった場合は、音声合成処理部4に、「はいかいいえで答えて?」という文章を表す単語データを供給する。音声合成処理部4は、この文章を読み上げる音声を表す音声データを生成して音声出力部5に供給し、音声出力部5にこの音声データが表す音声を再生させる。そして、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、いずれかのカテゴリに属する単語データが供給されると、当該単語データのカテゴリに従って、判別処理ステップQB1の処理における各遷移先のいずれかに処理を移す(判別処理ステップQB6)。

[0145] あるいは、エージェント処理部7は、トリガ取得処理ステップTG0では単語「つまない」を示す単語データが自然言語解析部3より供給されると、これを取得して判別処理ステップQB0に引き渡す。そして、判別処理ステップQB0では、この単語「つまない」が、他の単語との所属カテゴリの重なり合いなどの結果、どのカテゴリに属するものとして扱われるかを判別する。そして、カテゴリ「食事」に属すると判別すると、処理を判別処理ステップQB3に移し、他のカテゴリに属すると判別すると、該当するカテゴリに属する単語を取得するトリガ取得処理で当該単語データが取得されたものとして処理を継続する。

[0146] エージェント処理部7は、判別処理ステップQB2では、まず上述の判別処理ステップQB1と同様にして、音声合成処理部4を介し、音声出力部5に、「それじゃコンビニ行く?」という文章を読み上げる音声を再生させる。そして、自然言語解析部3から、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、該当する単語データが供給されると、これが「肯定」又は「否定」のいずれに属するかを判別す

る。そして、「肯定」に属すると判別すると入出力処理ステップSP2に処理を移し、「否定」に属すると判別すると入出力処理ステップSP3に処理を移す。また、これらのいずれのカテゴリの単語の単語データも供給されなかった場合は、音声出力部5に、「はいかいいえで答えて？」という文章を読み上げる音声を再生させ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、供給された単語データのカテゴリに従い、判別処理ステップQB2の処理における各遷移先のいずれかに処理を移す(判別処理ステップQB4)。

[0147] エージェント処理部7は、入出力処理ステップSP2では、音声合成処理部4を介し、音声出力部5に、「コンビニを案内します。」という文章を読み上げる音声を再生させ、処理を入出力処理ステップEX2に移す。そして、入出力処理ステップEX2では、ナビゲーション部61に、このエージェント装置が搭載されている車両の地図上における現在位置と、この地図上でのコンビニの位置を示す情報とを表示することを指示する制御信号を送る。ナビゲーション部61は、この指示に従い、車両の地図上における現在位置と、この地図上でのコンビニの位置を示す情報とを表示し、このコンビニへの案内を開始する。

[0148] 入出力処理ステップSP3でエージェント処理部7は、音声合成処理部4を介し、音声出力部5に、「安全運転に気をつけて。」という文章を読み上げる音声を再生させ、処理を入出力処理ステップEX3に移す。そして、入出力処理ステップEX3では、ナビゲーション部61に制御信号を送ることにより、ナビゲーション部61に、車両の地図上における現在位置を表示させる。

[0149] 一方、エージェント処理部7は、判別処理ステップQB3では、音声出力部5に、「お店は決まっていますか？」という文章を読み上げさせ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、該当する単語データが供給されると、いずれのカテゴリに属するかを判別する。そして、「肯定」に属すると判別すると判別処理ステップQB7に処理を移し、「否定」に属すると判別すると判別処理ステップBR1に処理を移す。また、これらのいずれのカテゴリの単語の単語データも供給されなかった場合は、音声出力部5に、「はいかいいえで答えて？」という文章を読み上げる音声を再生させ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待

機し、供給された単語データのカテゴリに従い、判別処理ステップQB3の処理における各遷移先のいずれかに処理を移す。

[0150] エージェント処理部7は、判別処理ステップQB7では、音声合成処理部4を介し、音声出力部5に、「お店の名前は？」という文章を読み上げさせて、単語データが供給されるのを待機し、供給されると、ナビゲーション部61に、この単語データが示すレストランを検索し、該当するレストランがあれば、車両の地図上における現在位置と、この地図上での当該レストランの位置を示す情報とを表示するよう指示する制御信号を送る。そして、該当するレストランがあれば、ナビゲーション部61が、車両の地図上における現在位置と、この地図上でのレストランの位置を示す情報とを表示し、このレストランへの案内を開始する(ステップEX100)。一方、該当するレストランがなければ、ナビゲーション部61がその旨の情報をエージェント処理部7に返すなどして、エージェント処理部7は処理を入出力処理ステップSP5に移す。入出力処理ステップSP5でエージェント処理部7は、音声出力部5に、「近くにはありません。」という文章を読み上げさせ、処理を判別処理ステップBR1に移す。

[0151] エージェント処理部7は、判別処理BR1では、クロック62より現在の時刻を示すデータを取得して、このデータが示す時刻が、(a) 12時より前、(b) 12時以降で14時より前、(c) 14時以降、のいずれであるかを判別する。そして、(a) 12時より前であると判別すると、音声出力部5に「モーニングですね。」という文章を読み上げさせ(入出力処理ステップSP8)、処理を判別処理ステップQB10に移す。(b) 12時以降で14時より前であると判別すると、音声出力部5に「ランチですね。」という文章を読み上げさせ(入出力処理ステップSP10)、処理を判別処理ステップQB12に移す。(c) 14時以降であると判別すると、音声出力部5に「お茶の時間ですね。」という文章を読み上げさせ(入出力処理ステップSP9)、処理を判別処理ステップQB10に移す。

[0152] 一方、エージェント処理部7は、判別処理ステップQB10では、音声出力部5に、「おすすめのお店を検索しますか？」という文章を読み上げさせ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、該当する単語データが供給されると、いずれのカテゴリに属するかを判別する。そして、「肯定」に属すると判別すると入出力処理ステップSP14に処理を移し、「否定」に属すると判別すると判別処理ス

テップQB12に処理を移す。また、これらのいずれのカテゴリの単語の単語データも供給されなかった場合は、音声出力部5に、「はいかいいえで答えて？」という文章を読み上げる音声を再生させ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、供給された単語データのカテゴリに従い、判別処理ステップQB10の処理における各遷移先のいずれかに処理を移す(判別処理ステップQB13)。

[0153] エージェント処理部7は、入出力処理ステップSP14では、ナビゲーション部61に、車両の現在位置近隣のレストランを検索し、索出したレストランの名称を示すデータを供給し、車両の地図上における現在位置と、この地図上での当該レストランの位置を示す情報とを表示するよう指示する制御信号を送る。そして、ナビゲーション部61がレストランの名称を供給すると、エージェント処理部7はこの名称を音声出力部5に読み上げさせる。一方、ナビゲーション部61は、車両の地図上における現在位置と、この地図上での当該レストランの位置を示す情報とを表示し、このレストランへの案内を開始する(ステップEX5)。

[0154] エージェント処理部7は、判別処理ステップQB12では、音声出力部5に、「過去経験DB(データベース)を検索しますか？」という文章を読み上げさせ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、該当する単語データが供給されると、いずれのカテゴリに属するかを判別する。そして、「肯定」に属すると判別すると入出力処理ステップSP15に処理を移し、「否定」に属すると判別すると判別処理ステップQB11に処理を移す。また、これらのいずれのカテゴリの単語の単語データも供給されなかった場合は、音声出力部5に、「はいかいいえで答えて？」という文章を読み上げる音声を再生させ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、供給された単語データのカテゴリに従い、判別処理ステップQB12の処理における各遷移先のいずれかに処理を移す(判別処理ステップQB14)。

[0155] エージェント処理部7は、入出力処理ステップSP15では過去経験データベースを、例えばナビゲーション部61が供給する情報が示す車両の現在位置を検索キーとして検索し、索出したレストランの名称を音声出力部5に読み上げさせる。そしてエージェント処理部7は、ナビゲーション部61に、車両の地図上における現在位置と、索出

したレストランのこの地図上での位置を示す情報とを表示するよう指示する制御信号を送る。ナビゲーション部61は、車両の地図上における現在位置と、この地図上での当該レストランの位置を示す情報とを表示し、このレストランへの案内を開始する(ステップEX6)。

[0156] エージェント処理部7は、判別処理ステップQB11では、音声出力部5に、「それじゃ和食・洋食・中華どれがいい?」という文章を読み上げさせ、単語「和食」、「洋食」又は「中華」を表す単語データが供給されるのを待機し、該当する単語データが供給されると、いずれの単語かを判別する。そして、「和食」と判別すると入出力処理ステップEX51に処理を移し、「洋食」と判別すると入出力処理ステップEX52に処理を移し、「中華」と判別すると入出力処理ステップEX53に処理を移す。また、これらのいずれの単語の単語データも供給されなかった場合は、音声出力部5に、「和食／洋食／中華で答えて?」という文章を読み上げる音声再生させ、単語「和食」、「洋食」又は「中華」を表す単語データが供給されるのを待機し、供給された単語データのカテゴリに従い、判別処理ステップQB11の処理における各遷移先のいずれかに処理を移す(判別処理ステップQB15)。

[0157] エージェント処理部7は、入出力処理ステップEX51では和食レストランデータベースを、例えばナビゲーション部61が供給する情報が示す車両の現在位置を検索キーとして検索し、処理を判別処理ステップQB16に移す。なお、エージェント処理部7は、トリガ取得処理ステップTG51で単語「和食」を表す単語データを取得したときも、入出力処理ステップEX51に処理を移すものとする。

また、入出力処理ステップEX52では洋食レストランデータベースを、入出力処理ステップEX53では中華レストランデータベースを、入出力処理ステップEX51と同様に検索し、図示するように、処理を判別処理ステップQB17又はQB18に移す。なお、エージェント処理部7は、トリガ取得処理ステップTG52で単語「洋食」を表す単語データを取得したときも入出力処理ステップEX52に処理を移すものとする。トリガ取得処理ステップTG53で単語「中華」を表す単語データを取得したときも入出力処理ステップEX53に処理を移すものとする。

[0158] 判別処理ステップQB16、QB17又はQB18で、エージェント処理部7は、入出力

処理ステップEX51、EX52又はEX53で索出したレストランの名称と、確認を促す文章とを、音声出力部5に読み上げさせて、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機する。そして、該当する単語データが供給されると、いずれのカテゴリに属するかを判別し、「肯定」に属すると判別すると入出力処理ステップSP17、SP18又はSP19に処理を移し、「否定」に属すると判別すると入出力処理ステップSP16に処理を移す。また、これらのいずれのカテゴリの単語の単語データも供給されなかった場合は、音声出力部5に、「はいかいいえで答えて？」という文章を読み上げる音声を再生させ、カテゴリ「肯定」又は「否定」に属する単語データが供給されるのを待機し、供給された単語データのカテゴリに従い、判別処理ステップQB16、QB17又はQB18の処理における各遷移先のいずれかに処理を移す(判別処理ステップQB19、QB20又はQB21)。

[0159] そしてエージェント処理部7は、入出力処理ステップSP17、SP18又はSP19では、音声出力部5に、「ご案内します。」という文章を読み上げさせ、ナビゲーション部61に、車両の地図上における現在位置と、入出力ステップEX51、EX52又はEX53で索出したレストランのこの地図上での位置を示す情報とを表示するよう指示する制御信号を送る。ナビゲーション部61は、車両の地図上における現在位置と、この地図上での当該レストランの位置を示す情報とを表示し、このレストランへの案内を開始する(ステップEX8、EX9又はEX10)。

[0160] 一方、入出力処理ステップSP16で、エージェント処理部7は、音声出力部5に、「わがままね。Webでレストラン情報を検索してみるね。」という文章を読み上げさせる。そして、エージェント処理部7は、自己の通信制御装置を介して外部のネットワークにアクセスし、現在位置を基準にしてレストランの位置を示す情報の検索を行う(入出力処理ステップEX7)。そして、該当する情報を索出すると、例えば、ナビゲーション部61に、車両の地図上における現在位置と、この地図上での当該レストランの位置を示す情報とを表示させる。

[0161] なお、上述の通り、エージェント処理部7に供給される単語データには音声認識部2によってスコアが付され、処理の遷移は、重み係数JとスコアSとの積 $S \cdot J$ が最大となるワイヤが示す内容に従って行われるので、フローがジャンプして上述した順序と異



なる処理が行われる場合があり得る。そして、トリガとなる処理項目を先行処理項目として有するワイヤに対しては、遷移定数 $k$ がある程度高い値に設定されているものとすれば、結果としてこのワイヤの重み係数 $J$ も高い値となり、ジャンプが起きやすくなって、エージェント処理部7は突然の話題転換にも対応できるようになる。例えば、「お腹減った」という音声入力を認識し、「食事に行く？」と質問した後に、「つままない」とユーザが発話した場合、単語「つままない」の単語データを取得する処理項目に対応するワイヤに設定されている単語「つままない」の認識スコア $S$ と重み係数 $J$ との積 $S \cdot J$ が算出され、また、判断処理「食事に行く？」に連なるワイヤに設定される単語のスコア $S$ と重み係数 $J$ との積 $S \cdot J$ も計算される(もちろん、他のワイヤに対してもそれぞれ計算される)。結果として、単語「つままない」について求めた積が最も大きな値を示せば、エージェント処理部7は処理のフローをジャンプし、単語「つままない」の単語データを取得する処理項目に後続する処理項目へと遷移する。

[0162] 以上説明したこのエージェント装置は、処理の内容を示すデータやワイヤが適切に記述されれば、ユーザの自然な発話に応答して、ユーザの欲求を満たすためにどの機器にどのような制御を加えればよいかを適切に判断して、判断結果に従った制御を機器に加えることができるようになる。

[0163] なお、このエージェント装置の構成は上述のものに限られない。

例えば、単語データベースD3は、必ずしも単語を示すデータのみならず、複数の単語からなる語句を示すデータを単語データベースD3の要素として記憶するようにしてもよいし、単語の一部あるいは音素を示すデータを単語データベースD3の要素として記憶するようにしてもよい。また、単語等は必ずしも特定の概念の下にグルーピングされている必要はなく、グルーピングを行う場合も、グルーピングを行うために用いられるデータは、必ずしもカテゴリ辞書の形をとっていなくてもよい。

[0164] また、音声認識部2は、ユーザが発話した音声のすべてを認識した結果を自然言語解析部3に供給する代わりに、ある特定の単語のみを認識して、認識した単語を直接にエージェント処理部7へと供給するようにしてもよい(ワードスポッティング方式)。この場合、このエージェント装置は、自然言語解析部3を備えている必要はない。

[0165] また、エージェント処理部7は、複数の単語が属するカテゴリを、これら複数の単語

がどのような組み合わせで同一の文章に含まれているかに基づいて決定してもよい。この場合、エージェント処理部7は、例えば図18にデータ構造を示すようなテーブルを記憶し、このテーブルを参照することにより、組をなす複数の単語のカテゴリを決定すればよい。なお、図18に例示するテーブルは、例えば、単語「お腹」と単語「減った」が同一の文章中に含まれていた場合は、これら2個の単語は、この文章に係る処理を行う限りではカテゴリ「食事」に属するものとして扱われることを示す。また、単語「設定」と単語「目的地」が同一の文章中に含まれていた場合は、これら2個の単語は、この文章に係る処理を行う限りではカテゴリ「ナビ設定」に属するものとして扱われることを示す。

- [0166] また、エージェント処理部7は、ワイヤに対応する遷移定数 $k$ を、過去に当該ワイヤが表す遷移を実行した数などに基づき所定の基準に従って変化させ、重み係数 $J$ が変化後の遷移定数 $k$ に基づいて計算された値となるようにワイヤを書き換えてもよい。

具体的には、例えば、ワイヤデータベースD5に、それぞれのワイヤについて、当該ワイヤが表す遷移が実行された回数を記憶しておく。そしてエージェント処理部7は、当該遷移が新たに行われる毎に、この回数の値を書き換えることにより、この回数の値を1ずつインクリメントし、それぞれのワイヤに対応する遷移定数 $k$ を、例えば、当該ワイヤについて記憶された回数に比例した値と書き換える。

なお、上述の実施例では、ワイヤに設定される条件は、それぞれのワイヤに対して設定されるとして説明したが、これには限られない。例えば、判別処理の中に各ワイヤの条件を記述するようにしてもよい。この場合、各条件がいずれのワイヤに対応するものかを予め特定しておく。

- [0167] また、エージェント処理部7は、判別処理や入出力処理において出力するデータを、これらの処理に引き渡されたデータや、これらの処理に伴って入力したデータや、その他任意の条件に従って変化させるようにしてもよい。

- [0168] また、このエージェント装置は、エージェント処理部7の制御に従って画像を出力するための表示装置(例えば、液晶ディスプレイ等)を備えていてもよく、エージェント処理部7は、入出力処理や判別処理において、処理毎に所定の画像を表示させるようこの表示装置を制御してもよい。

- [0169] また、エージェント処理部7は、トリガ取得処理、判別処理、入出力処理等の各種処理とワイヤとが全体として形成するフローを分担して行う、互いに接続された複数のデータ処理装置(例えば、コンピュータ等)から構成されていてもよい。この場合、エージェント処理部7を構成するそれぞれのデータ処理装置は、エージェント処理部7が実行し得るフロー全体のうち、自己が実行する可能性がある部分を表すデータを、処理項目データベースやワイヤデータベースの要素として記憶すれば十分である。そして、それぞれのデータ処理装置が記憶するデータが、当該データ処理装置が実行する部分の処理をマクロ定義するようなデータとなっていれば、複数のデータ処理装置に分散処理を行わせることも容易である。
- [0170] また、このエージェント装置は、音声入力部1や自然言語解析部3あるいは音声出力部5も複数備えていてよい。
- また、音声入力部1は、たとえば、音声を表すデータが記録された記録媒体(たとえば、フロッピー(登録商標)ディスクや、CD(Compact Disc)や、MO(Magneto-Optical Disk)など)から波形信号を読み出して自然言語解析部3に供給する記録媒体ドライブ装置(たとえば、フロッピー(登録商標)ディスクドライブや、CD-ROMドライブや、MOドライブなど)を備えていてもよい。
- [0171] また、このエージェント装置は、例えば、カーナビゲーションシステムの制御に限られず、他の機器の制御を行ってもよい。従って、車載機器の制御装置を制御してもよく、エアコン等の空調機器を制御してもよく、モーターの動力により開閉する窓の開閉や、ライトの点灯/消灯や、電動ワイパーの始動/停止を制御してもよい。
- [0172] また、オーディオ機器を制御してもよい。オーディオ機器の制御としては、例えば、ラジオのチューニングの制御、CD(Compact Disc)やMD(Mini Disc)、DVD(Digital Versatile Disc)等の記録媒体に記録された音声や画像を再生する記録媒体ドライブ装置の制御に適用可能である。
- [0173] オーディオ機器の制御の具体的な態様としては、例えば、CDプレーヤにより再生されている楽曲を変えたいと思ったユーザが指示のために用いる可能性のある表現(例えば、「曲を変えて」「別の曲」「違う曲」等)を構成する単語の組み合わせ(例えば、「曲」と「変えて」、「別」と「曲」、「違う」と「曲」、等)を、カテゴリ辞書を用いて「曲変更」

というカテゴリに分類しておき、CDプレーヤがCDに記録された楽曲の再生中に、カテゴリ「曲変更」に属する単語の組み合わせが発声され、該当する各単語を表す単語データが音声認識部2から自然言語解析部3を経てエージェント処理部7に供給されたとき、エージェント処理部7がこれに応答して(具体的には、カテゴリ「曲変更」に属する単語データを取得するトリガ処理項目の後続の処理項目に遷移して)、CDプレーヤに、再生中のトラックを変更することを指示する制御信号を送る、等が考えられる。尚、「曲」という単語と共に、「変えて」「別」「違う」の単語を1つのワイヤに条件として設定するようにしてもよい。これによって、自然言語解析部3から入力音声に対応する単語のスコアが直接ワイヤに対して設定される。自然言語解析部3から供給された単語データが、単語「曲」と単語「変えて」であった場合、それぞれのスコアが当該ワイヤに設定される。そして、この設定されたスコアに基づいて当該ワイヤの算出結果を求める。この場合、各スコアを合算してもよいし、または平均値を求めてもよい。そして、上述のカテゴリに各単語を分類する場合でも同様だが、ここで得られた各スコアとワイヤに設定される重み係数との乗算によってワイヤの算出結果が求められ、この算出結果が最も高い値を示すワイヤに状態を遷移することになる。

- [0174] また、他の具体例としては、エージェント処理部7が、音声出力部5に「この曲でいい？」という音声を出力させ、その後、「肯定」のカテゴリに属する単語(「うん」「いい」「OK」等)又は「否定」のカテゴリに属する単語(例えば「いや」「ちがう」「だめ」「別の」等)の単語データが供給されたとき、供給された単語データのカテゴリに従って、CDプレーヤに楽曲の再生や停止を指示することが考えられる。

#### [第2の実施の形態]

- [0175] 場合によって、ユーザはあいまいな表現を発することがある。例えば、「つまんない」という単語は、「音楽再生」「曲変更」「買い物」あるいは「食事」など、様々な要求を意図して発せられ得る。このようなあいまいな表現の指示が発せられたときにこの指示の意図を特定する本発明の第2の実施の形態のエージェント装置を説明する。
- [0176] このエージェント装置の物理的構成は、例えば、図1に示す第1の実施の形態の構成と実質的に同一である。ただし、このエージェント装置の入出力対象機器群6には、ナビゲーション部61やクロック62に加え、CDプレーヤが含まれているものとする。

また、カテゴリ辞書により定義されているカテゴリの少なくとも一部は、入出力対象機器群6に属する少なくともいずれかの機器と対応付けられているものとする。

[0177] このエージェント装置のエージェント処理部7に、単語「つまんない」を表す単語データが供給された場合、エージェント処理部7は、単語「つまんない」が属するカテゴリを検索する。その結果、例えば「音楽再生」「曲変更」「買い物」「食事」の4つのカテゴリに属していたとして、次に、エージェント処理部7は、入出力対象機器群6に属する機器のうち、これらのカテゴリに対応付けられている機器から、当該機器の現在の動作状態を示す情報を取得する。

[0178] 取得した情報が、CDプレーヤが何ら音声を再生していないことを示している場合、エージェント処理部7は、例えば「音楽でもかける？それとも買い物に行く？食事でいいく？」という音声を音声出力部5に発生させることにより、上述の4つのカテゴリのうち「曲変更」を除く「音楽再生」「買い物」「食事」の3つのカテゴリのいずれかを選択するよう促す。

[0179] そして、「CDをかけて」、あるいはその他、カテゴリ「CD再生」に属する単語ないしそのような単語を含む文章をユーザが発声し、カテゴリ「CD再生」に属する単語の単語データがエージェント処理部7に供給されると、エージェント処理部7は、カテゴリ「CD再生」に属する単語の単語データを取得するトリガ取得処理に後続する処理へと遷移する。この結果、例えば「CDを再生するね」という音声を音声出力部5に発生させ、CDプレーヤに音声の再生を指示する制御信号を送る。

[0180] また、単語「うた」を含む文章「なんか“うた”でも聞きたいなあ」、あるいはその他、カテゴリ「音楽」に属する単語ないしそのような単語を含む文章をユーザが発声し、カテゴリ「音楽」に属する単語の単語データがエージェント処理部7に供給されると、エージェント処理部7は、カテゴリ「音楽」に属する単語の単語データを取得するトリガ取得処理に後続する処理へと遷移する。

同様に、カテゴリ「食事」に属する単語ないしそのような単語を含む文章をユーザが発声し、カテゴリ「食事」に属する単語の単語データがエージェント処理部7に供給されると、エージェント処理部7は、カテゴリ「食事」に属する単語の単語データを取得するトリガ取得処理に後続する処理へと遷移する。この場合、例えば図11に示すQB

3に状態を遷移することとなる。

[0181] なお、入出力対象機器群6に属する機器の現在の動作状態を示す情報が、CDプレーヤが音声を再生していることを示している場合、エージェント処理部7は、例えば「曲を変える？それとも買い物に行く？食事にでもいく？」という音声を音声出力部5に発生させることにより、上述の4つのカテゴリのうち「音楽再生」を除く「曲変更」「買い物」「食事」の3つのカテゴリのいずれかを選択するよう促す。

[0182] 以上説明した動作を行うことによって、本発明の第2の実施の形態のエージェント装置は、ユーザのあいまいな指示を示す音声に基づいて、ユーザの意図を適切に特定する。

[第3の実施の形態]

[0183] 上述の各実施の形態のエージェント装置は、ユーザの要求を的確に判断し、また、あいまいな指示からもユーザの要求を具体的に導いて、機器の制御を行う。しかし、ユーザとしては、単に自らが一方的に発する要求を判断することだけでなく、話し相手を必要とする場合もあり得る。例えば、一人で車を運転するユーザは、そのような必要を感じる場合が多いと考えられる。しかし、車の運転中であれば、携帯電話を使用して会話をするのは危険が伴うし、また、通話料金がかさんでしまうという問題もある。

[0184] このような問題は、エージェント装置がユーザの話し相手になれば解決する。以下では、ユーザとの会話を達成する、本発明の第3の実施の形態に係るエージェント装置を説明する。

[0185] このエージェント装置の物理的構成は、例えば、上述した第1の実施の形態の構成と実質的に同一である。ただし、このエージェント装置の入出力対象機器群6は、FM多重放送を受信し文字情報を抽出してエージェント処理部7に随時供給するラジオ受信機、あるいは、外部の通信回線を介して文字情報を取得しエージェント処理部7に随時供給する端末を含んでいる受信機を含んでいるものとし、エージェント処理部7は、このラジオ受信機あるいは端末が供給する文字情報を自己の不揮発性メモリに蓄積するものとする。そして、蓄積した文字情報を読み出す際は、読み出しを行ったことを示すデータを生成して格納するものとする。

[0186] そして、例えば単語「なんかニュースない？」を表す単語データを取得するトリガ取

得処理が実行されると、このトリガ取得処理に後続する処理として、例えば以下説明する処理を実行する。

[0187] 例えば、エージェント処理部7は、自己の不揮発性メモリに蓄積された文字情報のうち、まだ読み出されていないものがあると、該当する文字情報のジャンルを知らせるメッセージ(例えば、「スポーツと、社会経済に関するニュースがあります」等)を読み上げる音声を音声出力部5に発生させる。

[0188] なお、該当する文字情報のジャンルは、受信された文字情報に予め分類を示す情報があればこれに基づいて特定する。ない場合には、エージェント処理部7が当該文字情報の内容に基づいて判断する。具体的には、例えば「〇〇党の幹部が汚職により更迭を余儀なくされ…」等の内容を含む文字情報では、「〇〇党」「汚職」の単語に基づいて、社会面記事と判断する。同様に、「〇〇株式会社が発表した決算発表によって同社の株価が高騰し…」との内容を含む文字情報では、「決算発表」「株価」の単語に基づいて、経済面記事と判断する。また「サッカーの〇〇(チーム名)が〇〇(いつの)試合で勝利し、優勝に王手をかけた」との内容を含む文字情報では、「サッカー」「試合」「優勝」の単語に基づいて、スポーツ面記事と判断する。

[0189] そして、ユーザが発した「スポーツのニュースを教えて」という文章を表す単語データの集合が供給されると、エージェント処理部7は、自ら蓄積している文字情報のうち、まだ読み出されていないスポーツ面記事を反映した文章(例えば、「サッカーの〇〇(チーム名)が〇〇(いつの)試合で勝利し、優勝に王手をかけたんだって」等)を読み上げる音声を音声出力部5に発生させる。

[0190] なお、ニュース等で受信する文字情報は文末に「ーした」等の表現が使われることが多く、通常の会話にはそぐわない。そこで、文字情報をそのまま読み上げるようにせず、上述のように、文章の語尾を「んだって」等、友人同士での会話の語尾に使われそうな言葉に置換して読み上げるようにするとよい。

[0191] また、エージェント処理部7は、音声出力部5に文字情報を読み上げさせる際、文字情報の一部(例えば、冒頭の段落)のみをまず読み上げさせ、ユーザが当該文字情報の詳細を求める音声(例えば、「もっと詳しい情報ある?」と問い掛ける音声)を発するのを待機し、発したとき、これに応答して残りの部分を読み上げさせるようにし

てもよい。

[0192] また、エージェント処理部7は、ユーザが当該文字情報の詳細を求める音声を発したことに応答して、自ら蓄積している文字情報を検索し、既に全文を読み上げたか否かを判別して、読み上げたと判別した場合には、「ううん、今のところこれだけだね」という音声を出力させ、また、詳細部分がある場合には「あるよ、じゃあ全文読み上げるね」という音声を出力させた後に、該当する詳細部分を読み上げる音声を出力させるようにしてもよい。

[0193] また、エージェント処理部7は、音声出力部5に読み上げさせた文字情報に対する質問を含んだ音声をユーザが発したとき、これに応答して、自らが蓄積している文字情報のうちから、当該質問への回答となり得るものを索出して音声出力部5に読み上げさせてもよい。例えばユーザが、「どことの試合で勝ったの?」という質問を発した場合、当該試合の結果を解説する文字情報を検索し、索出されれば、この文字情報を読み上げさせる、等である。また、「次の試合はいつ?」と質問された場合も、次の試合のスケジュールを解説する文字情報を検索し、索出されればこれを読み上げさせる等すればよい。

[0194] そして、その結果ユーザが更に、「チケット取れる?」等と質問すると、エージェント処理部7は、これまでのユーザとの会話内容に基づいて「チーム名」や「試合日」を特定し、例えば「ちょっと調べてみるね」という音声を音声出力部5に発生させた後に、外部のネットワーク等を通じてチケット販売サイトに接続し、該当するチケットの有無を照会する。

[0195] その結果、当該チケットがまだ購入可能であるとの回答が得られた場合、エージェント処理部7は、「今ならまだ大丈夫。○席で○○円だけど、購入する?」などとチケットの販売状況を説明する音声を音声出力部5に読み上げさせることにより、ユーザの回答を促す。これに対してユーザが「購入する」「注文して」「買う」等、肯定を意味する単語を発声した場合、この単語を表す単語データに応答して、チケットの購入を支援する処理(例えば、ネット決済等を支援する公知の処理)へと移行する。

[0196] 一方、チケットが既に完売になっているとの回答が得られた場合、エージェント処理部7は、「残念だけど、もう満席みたい。」等の文章を音声出力部5に読み上げさせて



、一連の処理を終了し、その他に、まだ未出力の文字情報があれば、該当する文字情報を出力する処理に移行する。

[0197] 以上説明した動作を行うことによって、本発明の第3の実施の形態のエージェント装置は、単にユーザの所望する制御を行うのみでなく、ユーザとの会話を実現する。尚、上述のエージェント装置とユーザの会話は、オーディオ装置に対しても適用することが可能である。例えば、オーディオ装置は、インターネット等を通じて予めユーザが登録したアーティスト条件に基づいて情報を取得してメモリする。取得された情報は、アーティスト条件に基づいて分類される。そして、ユーザが「新着情報ない？」等を問い掛けると、エージェント装置はメモリ内容を検索し、まだ提示していない情報があれば、音声出力によってこれを提示する。もちろんこのとき、映像情報を伴うものであってもよい。例えばこの新着情報が新しいアルバムに関する情報であれば、更にエージェント装置はこのアルバムの購入をユーザに確認し、了解が取れるとこれを購入するようにインターネットCDショップに購入手続きをとる。また同様に、コンサート情報であれば、エージェント装置はこのコンサートのチケットの購入をユーザに確認し、了解が取れるとこれを購入するようにインターネットチケットショップに購入手続きをとる。尚、これらの音声認識処理等は上述の実施例に基づいて行われる。

[0198] 以上、この発明の実施の形態を説明したが、上述の実施の形態に示されるような以下のいずれかの処理を、又は組み合わせて施すことによって、ユーザが発話する様々な表現の音声を適切に認識して、対応する処理を実行することが可能になる。

1. 各ワイヤに対して、当該ワイヤにおいて同じ意味として取り扱われる単語を複数条件として設定する。
2. 各ワイヤに対して、一若しくは複数の単語が分類されるカテゴリを設定する。
3. 各ワイヤに対して、異なる意味の単語若しくはカテゴリを複数条件として設定する。
4. 複数の単語が同一の文章(入力音声)に含まれていた場合に、これらの単語が共通して属するカテゴリのみを、これらの単語が属するカテゴリとして扱うことにより、これらの単語が発話された意図を文脈に応じて把握することを可能とする。
5. 単語データが有するカテゴリを、当該単語データの意味として扱う。

6. トリガ取得処理で取得されるデータが単語データである場合は、ワイヤに設定される条件は、当該単語データが表す単語に代えて、当該単語データが表す単語がグルーピングされているカテゴリを記述する。

7. 入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促すようにする。

8. 外部から取得した情報に基づいて、ユーザに提示する内容を変更する。

尚、この発明にかかる機器制御装置は、専用のシステムによらず、通常のコンピュータシステムを用いて実現可能である。

例えば、入出力対象機器群6に接続されたパーソナルコンピュータに上述の音声入力部1、自然言語解析部3、音声合成処理部4、音声出力部5及びエージェント処理部7の動作を実行させるためのプログラムを格納した記録媒体から該プログラムをインストールすることにより、上述の処理を実行するエージェント装置を構成することができる。そして、このプログラムを実行するパーソナルコンピュータが、図1のエージェント装置の動作に相当する処理として、例えば、図9(a)に示すフローを実行するものとする。

[0199] なお、パーソナルコンピュータに上述のエージェント装置の機能を行わせるプログラムは、例えば、通信回線の掲示板(BBS)にアップロードし、これを通信回線を介して配信してもよく、また、このプログラムを表す信号により搬送波を変調し、得られた変調波を伝送し、この変調波を受信した装置が変調波を復調してこのプログラムを復元するようにしてもよい。そして、このプログラムを起動し、OSの制御下に、他のアプリケーションプログラムと同様に実行することにより、上述の処理を実行することができる。

[0200] なお、OSが処理の一部を分担する場合、あるいは、OSが本願発明の1つの構成要素の一部を構成するような場合には、記録媒体には、その部分を除いたプログラムを格納してもよい。この場合も、この発明では、その記録媒体には、コンピュータが実行する各機能又はステップを実行するためのプログラムが格納されているものとする。

[0201] 本発明は、2003年12月26日に出願された、特願2003-436976に基づき、その明細書、特許請求の範囲、図面および要約書を含む。上記出願における開示は、本

明細書中にその全体が参照として含まれる。

#### 産業上の利用可能性

[0202] 本発明は、人間が言語の形で発する指示に適切に応答して機器を制御できる種々の機器制御装置等に利用可能である。

## 請求の範囲

- [1] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする機器制御装置。
- [2] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項1に記載の機器制御装置。
- [3] 前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された複数の語句の組み合わせが所定の条件を満たしているか否かを判別し、判別結果に基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項2に記載の機器制御装置。
- [4] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項1に記載の機器制御装置。
- [5] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項1に記載の機器制御装置。

- [6] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定することを特徴とする請求項1に記載の機器制御装置。
- [7] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行する、  
ことを特徴とする請求項1に記載の機器制御装置。
- [8] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促す、  
ことを特徴とする請求項1に記載の機器制御装置。
- [9] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択する、  
ことを特徴とする請求項1に記載の機器制御装置。
- [10] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を

出力する、

ことを特徴とする機器制御装置。

- [11] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とする音声認識装置。

- [12] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、

前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項11に記載の音声認識装置。

- [13] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項11に記載の音声認識装置。

- [14] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項11に記載の音声認識装置。

- [15] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定する

ことを特徴とする請求項11に記載の音声認識装置。

- [16] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行する、

ことを特徴とする請求項11に記載の音声認識装置。

- [17] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促す、

ことを特徴とする請求項11に記載の音声認識装置。

- [18] 外部機器からの情報を取得する情報取得手段(7)を更に備え、

前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択する、

ことを特徴とする請求項11に記載の音声認識装置。

- [19] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とする音声認識装置。

- [20] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とするエージェント装置。

- [21] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、

前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項20に記載のエージェント装置。

- [22] 前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された複数の語句の組み合わせが所定の条件を満たしているか否かを判別し、判別結果に基づいて、前記音声の発話者の発話内容を特定する、
- ことを特徴とする請求項21に記載のエージェント装置。

- [23] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定する、
- ことを特徴とする請求項20に記載のエージェント装置。

- [24] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定する、
- ことを特徴とする請求項20に記載のエージェント装置。

- [25] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定する、



ことを特徴とする請求項20に記載のエージェント装置。

- [26] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行する、

ことを特徴とする請求項20に記載のエージェント装置。

- [27] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促す、

ことを特徴とする請求項20に記載のエージェント装置。

- [28] 外部機器からの情報を取得する情報取得手段(7)を更に備え、

前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択する、

ことを特徴とする請求項20に記載のエージェント装置。

- [29] 前記処理実行手段(7)は、実行すべき処理として特定した処理が、外部から受信した情報を発話者に提示する処理であるとき、当該情報を読み上げる音声を発生させることにより当該提示を実行する手段を備える、

ことを特徴とする請求項20に記載のエージェント装置。

- [30] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を

出力する、  
ことを特徴とするエージェント装置。

- [31] 外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする車載機器制御装置。

- [32] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [33] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [34] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [35] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており

、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [36] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行する、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [37] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促す、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [38] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択する、  
ことを特徴とする請求項31に記載の車載機器制御装置。

- [39] 外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情

報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とする車載機器制御装置。

- [40] 車両に搭載可能に構成されたナビゲーション装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定して、当該ナビゲーション処理を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とするナビゲーション装置。

- [41] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、  
前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項40に記載のナビゲーション装置。

- [42] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項40に記載のナビゲーション装置。

- [43] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項40に記載のナビゲーション装置。

- [44] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており

、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴリに基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項40に記載のナビゲーション装置。

[45] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行する、  
ことを特徴とする請求項40に記載のナビゲーション装置。

[46] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促す、  
ことを特徴とする請求項40に記載のナビゲーション装置。

[47] 外部機器からの情報を取得する情報取得手段(7)を更に備え、  
前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択する、  
ことを特徴とする請求項40に記載のナビゲーション装置。

[48] 車両に搭載可能に構成されたナビゲーション装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定する処理特定手段(7)と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、  
を備え、  
前記処理特定手段(7)が特定したナビゲーション処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段は当該情報に基づ

いて音声を出力する、  
ことを特徴とするナビゲーション装置。

- [49] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべき音声処理の内容を特定して、当該音声処理を実行し、又は外部の機器に実行させるよう当該外部の機器を制御する処理実行手段(7)と、

を備える、

ことを特徴とするオーディオ装置。

- [50] 前記音声認識手段(2, 3)は、特定した語句の品詞を特定する品詞特定手段(3)を備え、

前記特定手段(7)は、前記音声認識手段(2, 3)が特定した語句のうち、所定の品詞であると特定された語句のみに基づいて、前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項49に記載のオーディオ装置。

- [51] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した語句が分類されるカテゴリに基づいて前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項49に記載のオーディオ装置。

- [52] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、異なる意味の語句若しくはカテゴリを複数個対応付ける対応情報を保持しており、前記音声認識手段(2, 3)が特定した語句若しくはカテゴリの組み合わせと前記対応情報に基づいて前記音声の発話者の発話内容を特定する、

ことを特徴とする請求項49に記載のオーディオ装置。

- [53] 前記特定手段(7)は、語句を1個以上のカテゴリに対応付ける情報を保持しており、前記音声認識手段(2, 3)が特定した複数の語句が共通して分類されているカテゴ

りに基づいて、前記音声の発話者の発話内容を特定する、  
ことを特徴とする請求項49に記載のオーディオ装置。

- [54] 前記特定手段(7)は、前記処理実行手段(7)のそれぞれの処理に対して、複数の語句を割り当てて保持しており、前記音声認識手段(2, 3)が特定した語句のうち少なくとも一つが前記処理に対して割り当てられた語句であるとき、対応する処理を実行する、

ことを特徴とする請求項49に記載のオーディオ装置。

- [55] 前記特定手段(7)は、入力された音声の意味が判別できなかったとき、より判別しやすい表現での入力を促す、

ことを特徴とする請求項49に記載のオーディオ装置。

- [56] 外部機器からの情報を取得する情報取得手段(7)を更に備え、

前記特定手段(7)は、前記情報取得手段(7)によって得られた情報に基づいて、出力する出力内容を選択する、

ことを特徴とする請求項49に記載のオーディオ装置。

- [57] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段と、

特定した内容に基づいて、実行すべき音声処理の内容を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した音声処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声出力する、

ことを特徴とするオーディオ装置。

- [58] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行ステップと、  
より構成される、  
ことを特徴とする機器制御方法。
- [59] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、より構成され、  
前記処理特定ステップで特定した制御が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声を出力する、  
ことを特徴とする機器制御方法。
- [60] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行ステップと、より構成される、  
ことを特徴とする音声認識方法。



- [61] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて実行すべき処理を特定する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、  
より構成され、  
前記処理特定ステップで特定した処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声を出力する、  
ことを特徴とする音声認識方法。
- [62] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行ステップと、  
より構成される、  
ことを特徴とするエージェント処理方法。
- [63] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて実行すべき処理を特定する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステッ

ブと、

より構成され、

前記処理特定ステップで特定した処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声を出力する、

ことを特徴とするエージェント処理方法。

[64] 車両に搭載された車載機器を制御するための車載機器制御方法であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行ステップと、

より構成される、

ことを特徴とする車載機器制御方法。

[65] 車両に搭載された車載機器を制御するための車載機器制御方法であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定する処理特定ステップと、

所定の通信装置を介して情報を取得する情報取得ステップと、

前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、

より構成され、

前記処理特定ステップが特定した制御が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップは当該情報に基づいて音声を

出力する、

ことを特徴とする車載機器制御方法。

- [66] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定して、当該ナビゲーション処理を実行する処理実行ステップと、

より構成される、

ことを特徴とするナビゲーション方法。

- [67] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定する処理特定ステップと、

所定の通信装置を介して情報を取得する情報取得ステップと、

前記情報取得ステップで取得した情報に基づいて音声出力する音声出力ステップと、

より構成され、

前記処理特定ステップで特定したナビゲーション処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップでは当該情報に基づいて音声出力する、

ことを特徴とするナビゲーション方法。

- [68] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、

前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、

特定した内容に基づいて、実行すべき音声処理の内容を特定して、当該音声処理を実行し、又は外部のオーディオ装置に実行させるよう当該外部のオーディオ装置を制御する処理実行ステップと、  
より構成される、  
ことを特徴とするオーディオ装置制御方法。

- [69] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識ステップと、  
前記音声認識ステップで特定した語句に基づき、前記音声の発話者の発話内容を特定する特定ステップと、  
特定した内容に基づいて、外部のオーディオ装置に実行させるべき音声処理の内容を特定する処理特定ステップと、  
所定の通信装置を介して情報を取得する情報取得ステップと、  
前記情報取得ステップで取得した情報に基づいて音声を出力する音声出力ステップと、  
より構成され、  
前記処理特定ステップで特定した音声処理が、前記情報取得ステップで取得された情報を出力するものであるとき、前記音声出力ステップで当該情報に基づいて音声を出力する、  
ことを特徴とするオーディオ装置制御方法。

- [70] コンピュータを、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、  
ことを特徴とする機器制御装置として機能させるためのプログラム。

[71]        コンピュータを、  
              音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声  
              が表す語句を特定する音声認識手段(2, 3)と、  
              前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する  
              特定手段と、  
              特定した内容に基づいて、制御の対象である外部の機器に対して実行すべき制御の内容を  
              特定する処理特定手段(7)と、  
              所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
              前記情報取得手段(7)が取得した情報に基づいて音声を入力する音声入力手段(5)と、  
              を備え、  
              前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を入力する  
              ものであるとき、前記音声入力手段(5)は当該情報に基づいて音声を入力する、  
              ことを特徴とする機器制御装置として機能させるためのプログラム。

[72]        コンピュータを、  
              音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該  
              音声を表す語句を特定する音声認識手段(2, 3)と、  
              前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する  
              特定手段と、  
              特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行  
              手段(7)と、  
              を備える、  
              ことを特徴とする音声認識装置として機能させるためのプログラム。

[73]        コンピュータを、  
              音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該  
              音声を表す語句を特定する音声認識手段(2, 3)と、  
              前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内

容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とする音声認識装置として機能させるためのプログラム。

[74] コンピュータを、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定して、当該処理を実行する処理実行手段(7)と、

を備える、

ことを特徴とするエージェント装置として機能させるためのプログラム。

[75] コンピュータを、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて実行すべき処理を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、  
ことを特徴とするエージェント装置として機能させるためのプログラム。

[76] コンピュータを、

外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定して、当該制御を実行する処理実行手段(7)と、  
を備える、

ことを特徴とする車載機器制御装置として機能させるためのプログラム。

[77] コンピュータを、

外部の車載機器を搭載した車両に搭載可能に構成された車載機器制御装置であって、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、前記車載機器に対して実行すべき制御の内容を特定する処理特定手段と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、

を備え、

前記処理特定手段(7)が特定した制御が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、  
ことを特徴とする車載機器制御装置として機能させるためのプログラム。

[78] コンピュータを、

車両に搭載可能に構成されたナビゲーション装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定して、当該ナビゲーション処理を実行する処理実行手段(7)と、  
を備える、

ことを特徴とするナビゲーション装置として機能させるためのプログラム。

[79] コンピュータを、

車両に搭載可能に構成されたナビゲーション装置であって、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段と、  
前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段と、  
特定した内容に基づいて、実行すべきナビゲーション処理の内容を特定する処理特定手段と、  
所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、  
前記情報取得手段(7)が取得した情報に基づいて音声を出力する音声出力手段(5)と、  
を備え、

前記処理特定手段(7)が特定したナビゲーション処理が、前記情報取得手段(7)



で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声を出力する、

ことを特徴とするナビゲーション装置として機能させるためのプログラム。

[80] コンピュータを、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべき音声処理の内容を特定して、当該音声処理を実行し、又は外部の機器に実行させるよう当該外部の機器を制御する処理実行手段(7)と、

を備える、

ことを特徴とするオーディオ装置として機能させるためのプログラム。

[81] コンピュータを、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声を表す語句を特定する音声認識手段(2, 3)と、

前記音声認識手段(2, 3)が特定した語句に基づき、前記音声の発話者の発話内容を特定する特定手段(7)と、

特定した内容に基づいて、実行すべき音声処理の内容を特定する処理特定手段(7)と、

所定の通信手段(74)を介して情報を取得する情報取得手段(7)と、

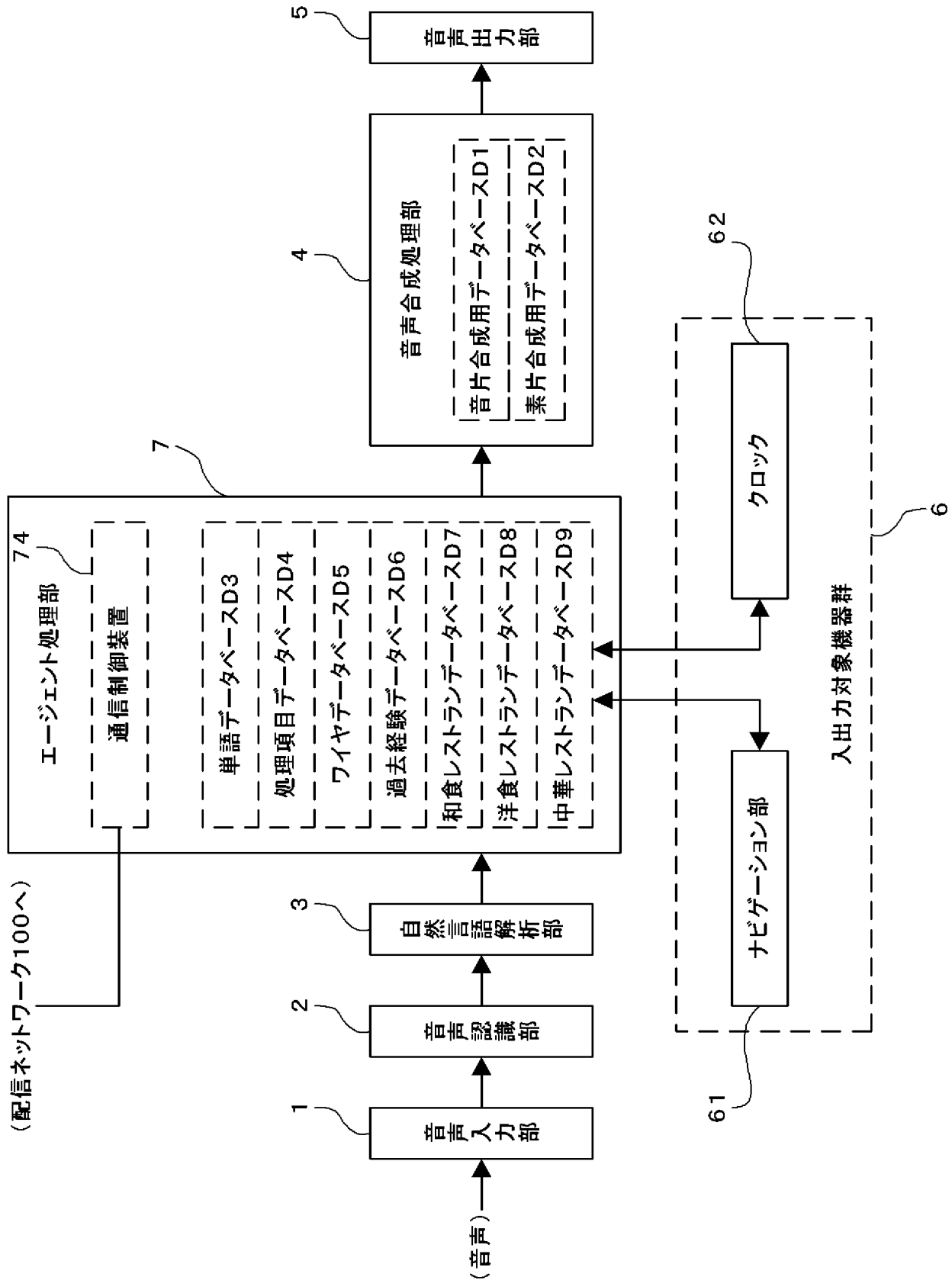
前記情報取得手段(7)が取得した情報に基づいて音声出力する音声出力手段(5)と、

を備え、

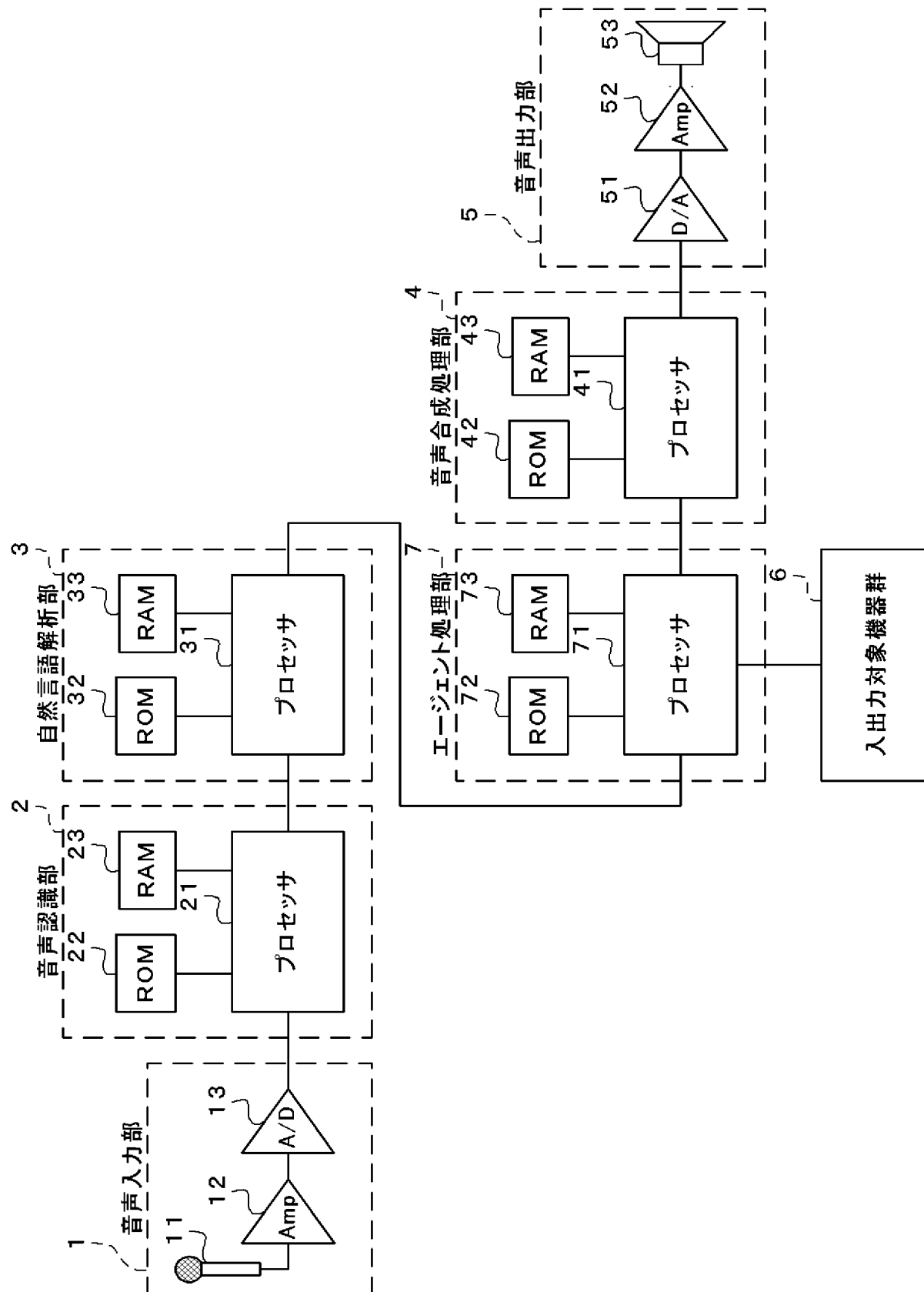
前記処理特定手段(7)が特定した音声処理が、前記情報取得手段(7)で取得された情報を出力するものであるとき、前記音声出力手段(5)は当該情報に基づいて音声出力する、

ことを特徴とするオーディオ装置として機能させるためのプログラム。

[図1]



[図2]



[図3]

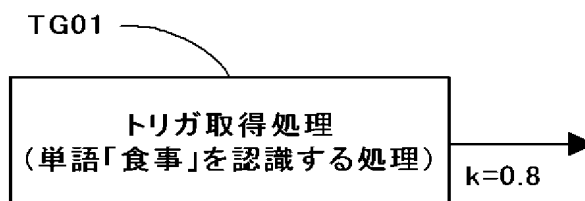
認識した単語	カテゴリ			
	食事	病院	旅行	ナビ設定
「減った」	○			
「すいた」	○			
「お腹」	○	○		
「食べ」	○			
「行く」	○	○	○	○
「設定」	○	○		○
「どこか」	○	○	○	○
「どこに」	○			
「食事」	○			
「痛い」		○		
「病院」		○		
「目的地」				○

[図4]

(a)

トリガ取得処理 遷移定数k	
記号: TG	
TG00: 「つまんない」	0.7
TG01: 「食事」	0.8
TG51: 「和食」	0.5
TG52: 「洋食」	0.5
TG53: 「中華」	0.5
⋮	⋮

(b)

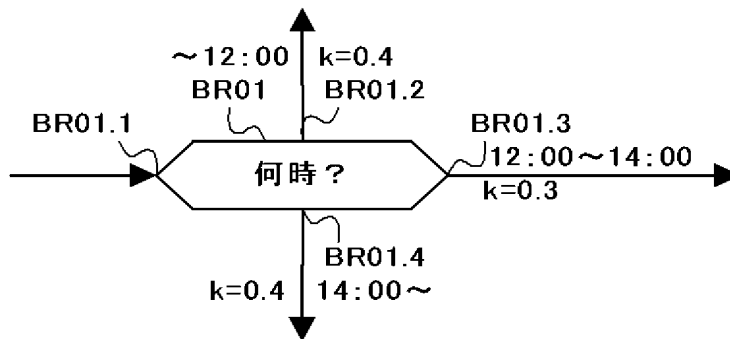


[図5]

(a)

判別処理	遷移定数k
記号:BRxx	
BR01:何時?	~12:00 :0.4
	12:00~14:00 :0.3
	14:00~ :0.4

(b)

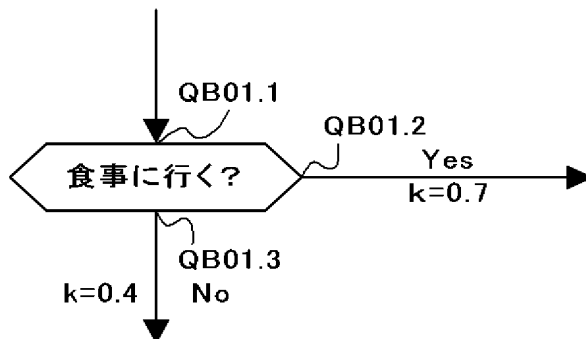


[図6]

(a)

判別処理(問い合わせあり)	遷移定数k
記号:QBxx	
QB01:食事に行く?	Yes:0.7 No:0.4
QB02:それじゃコンビニに行く?	Yes:0.5 No:0.3
⋮	⋮

(b)

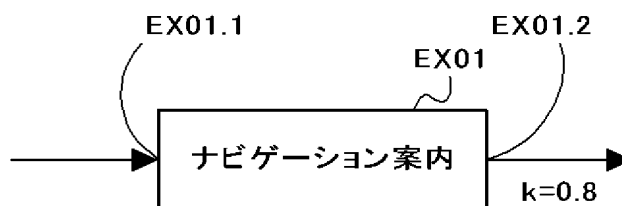


[図7]

(a)

出力処理	遷移定数k
記号: EXxx, SPxx	
EX02: ナビゲーション案内	0.8
⋮	⋮
SP02: コンビニ店案内	0.8
SP03: 安全運転に気を付けて	0.8
⋮	⋮

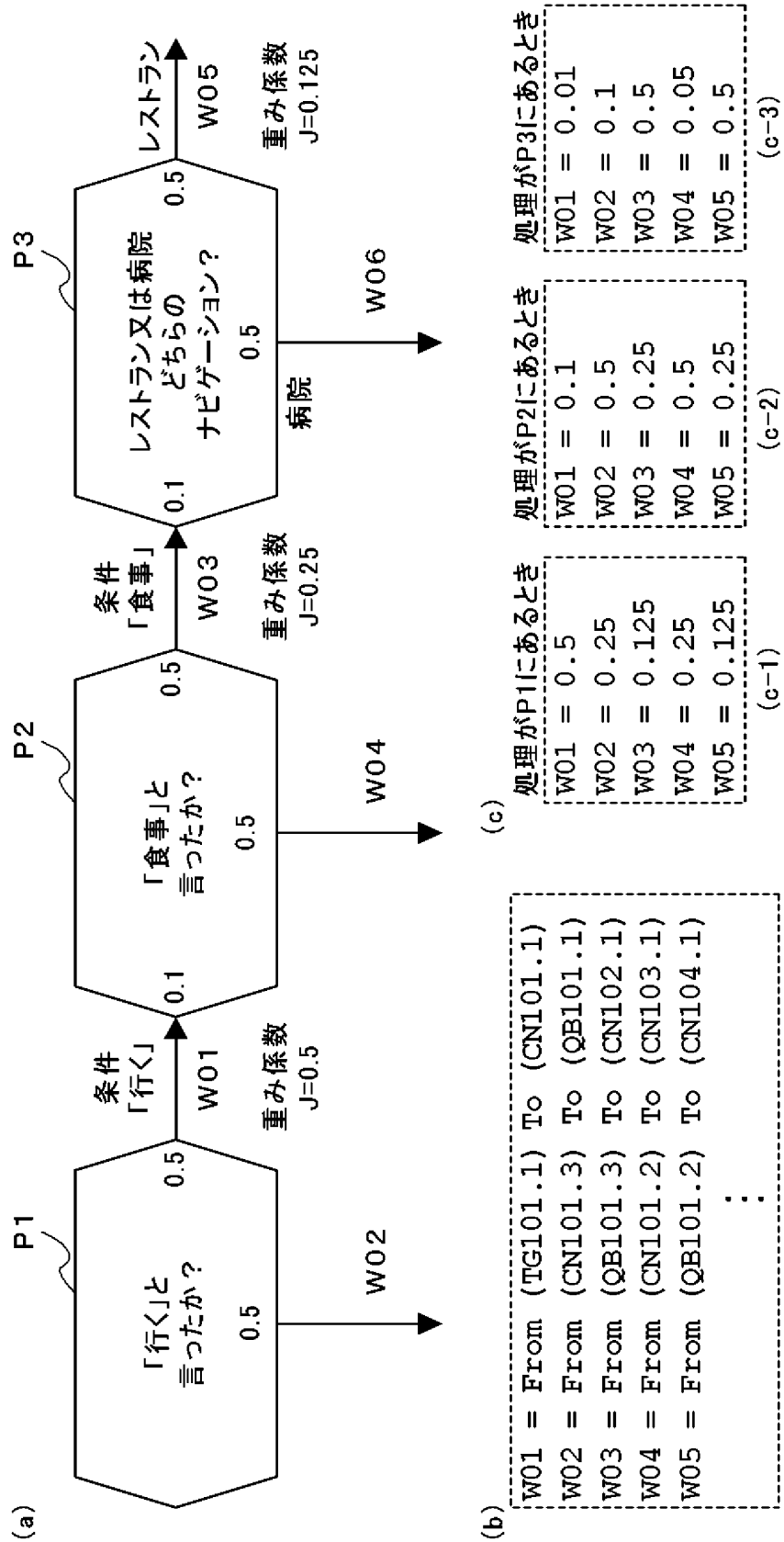
(b)



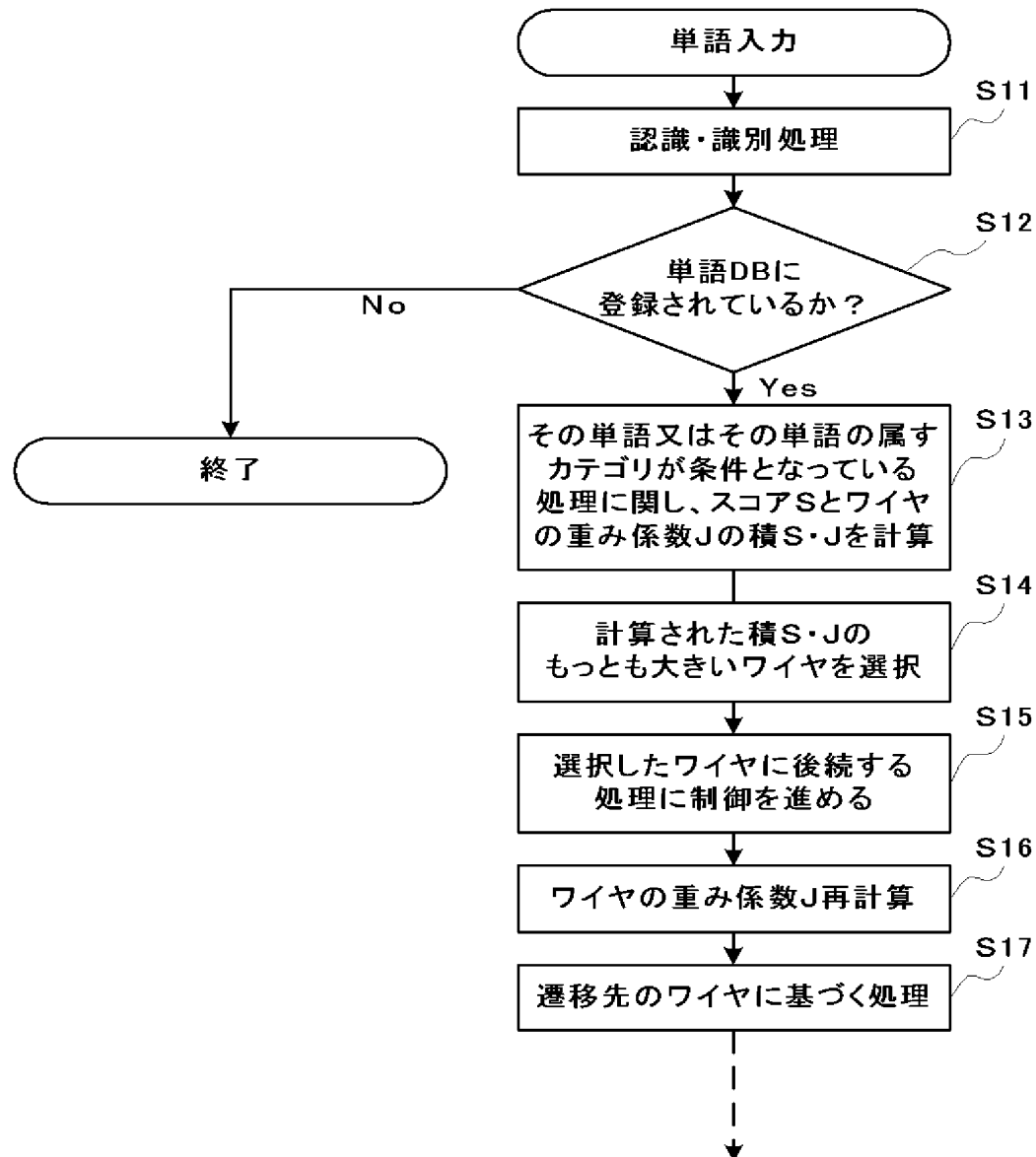
[図8]

ワイヤ番号Wn		先行ノード	後行ノード	重み係数J
W1	=	From (TG01.1)	To (CN01.1) ,	1.0
W2	=	From (CN01.3)	To (EX01.1) ,	0.5
W3	=	From (CN01.2)	To (QB01.1) ,	0.5
W4	=	From (TG02.2)	To (EX03.1) ,	1.0
W5	=	From (QB01.3)	To (EX03.1) ,	0.5
W6	=	From (QB01.2)	To (EX02.1) ,	0.5
W7	=	From (EX02.2)	To (TG03.3) ,	0.5
		⋮		

[図9]

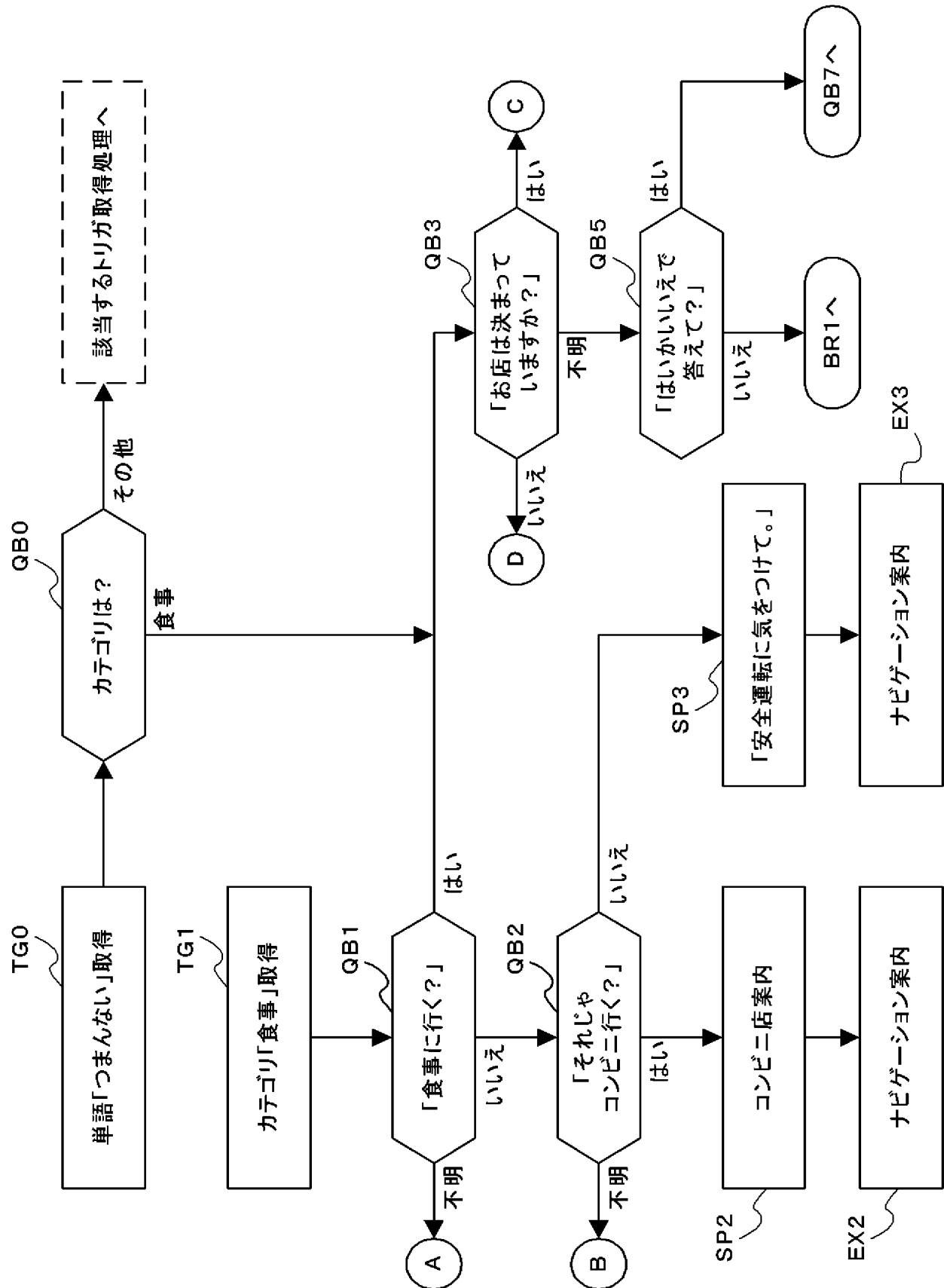


[図10]

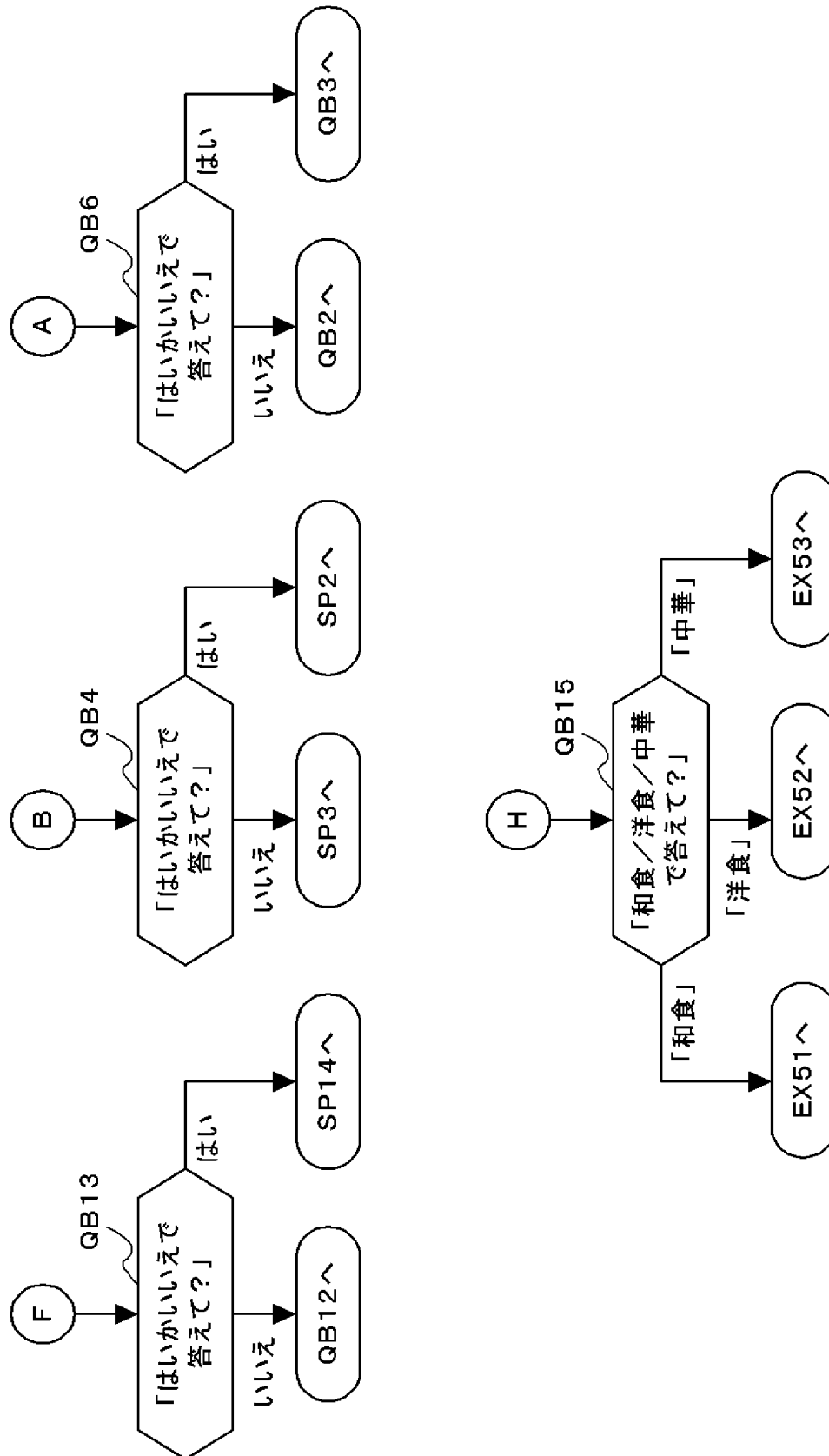


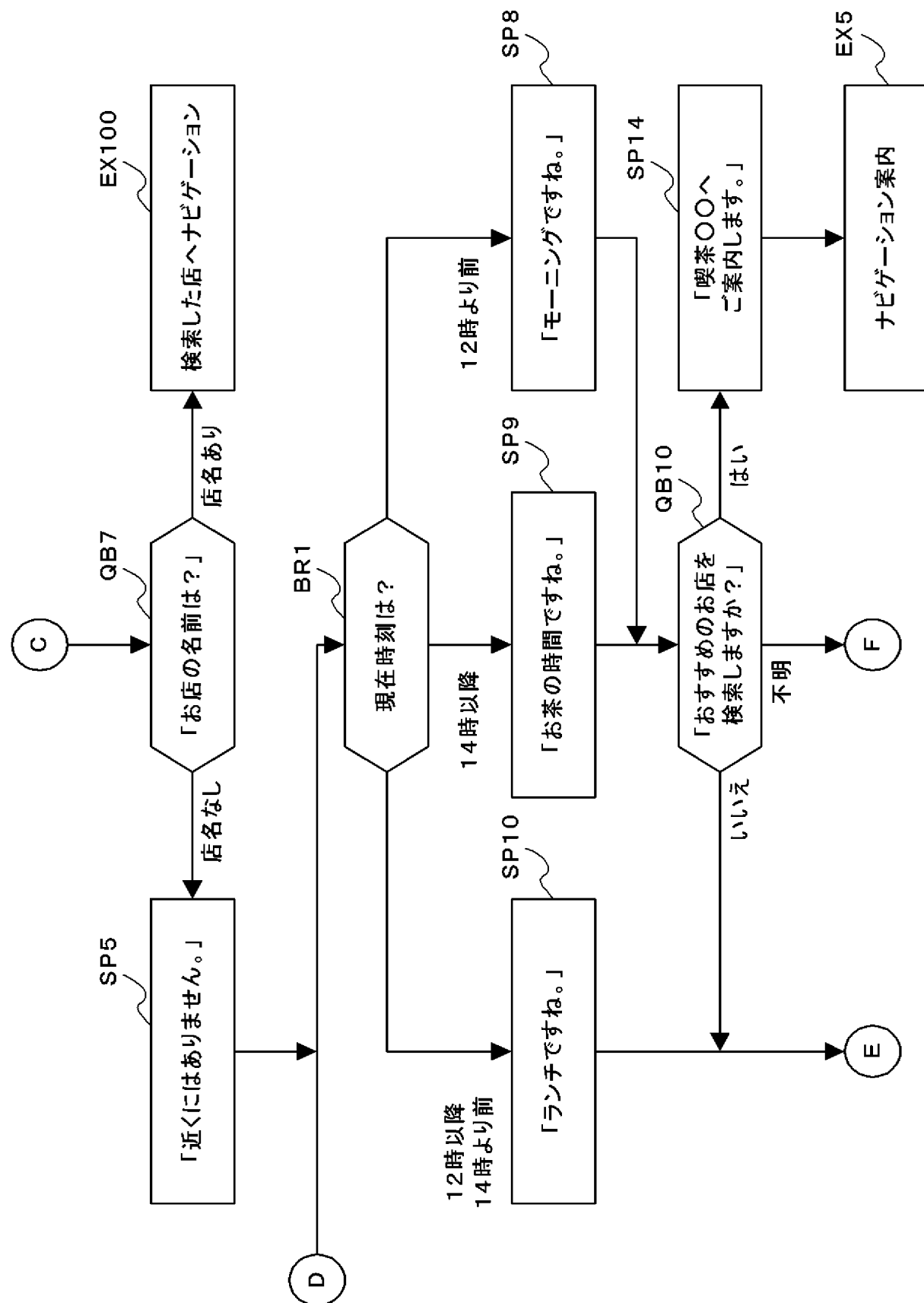


[図11]

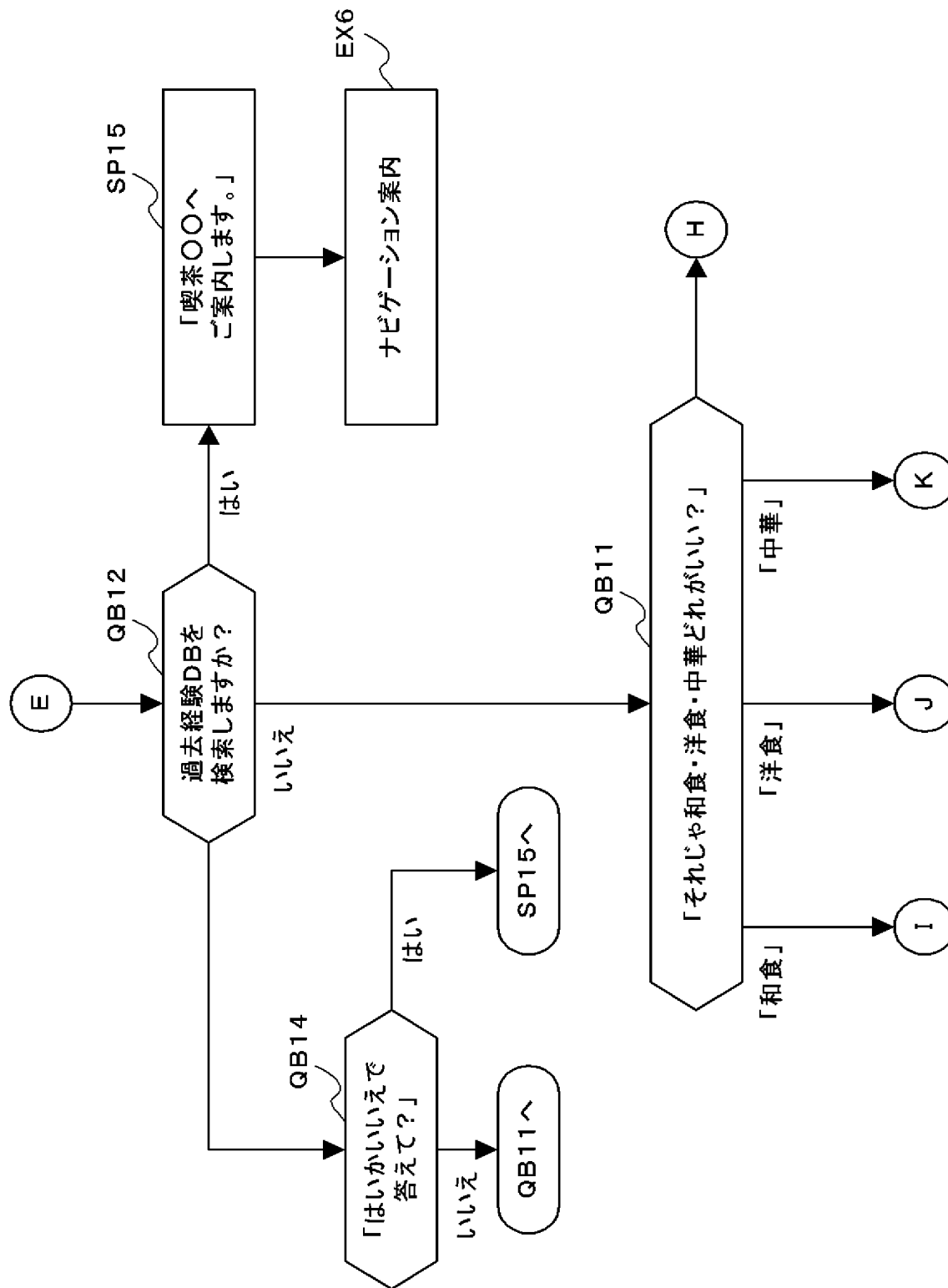


[図12]

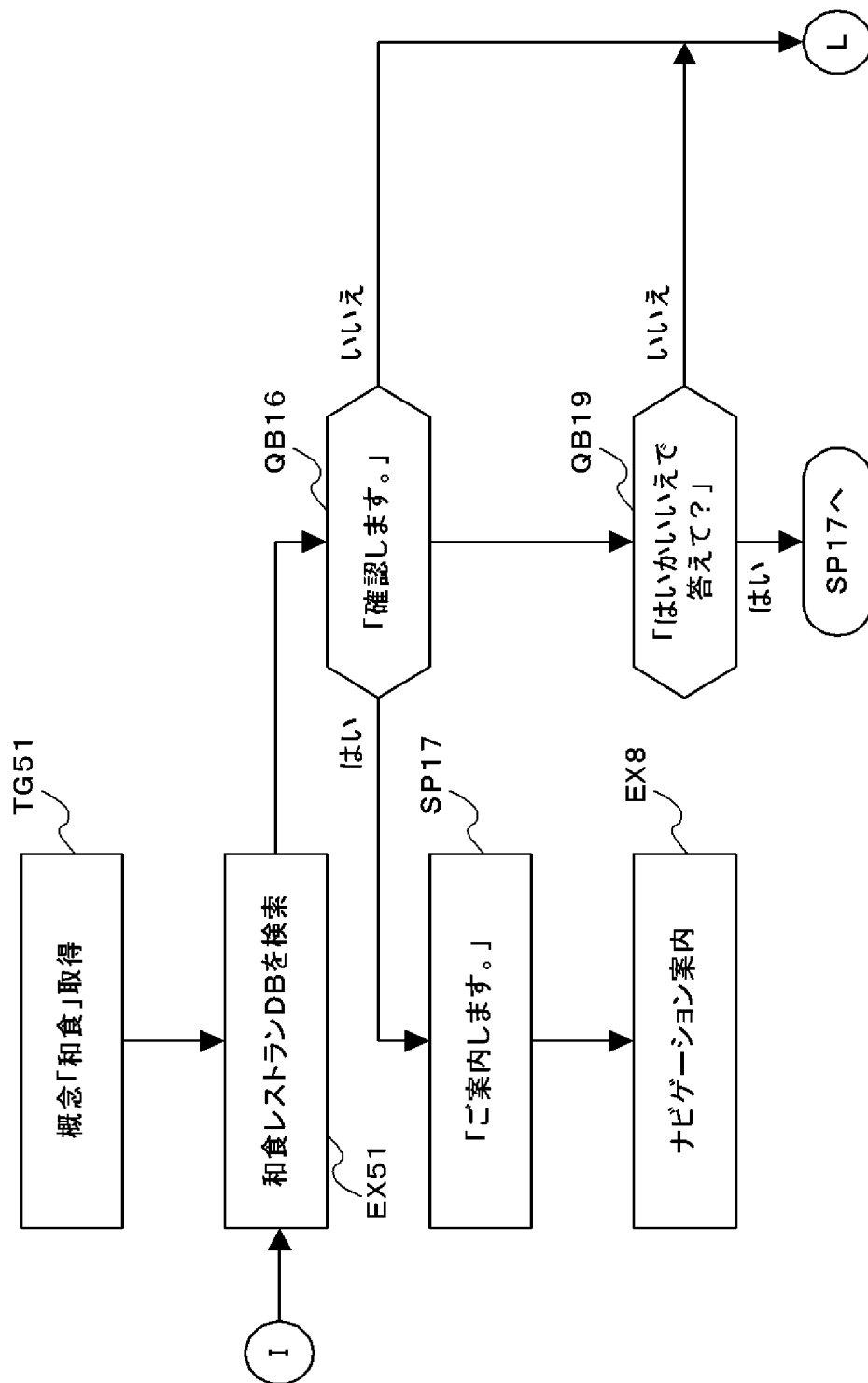




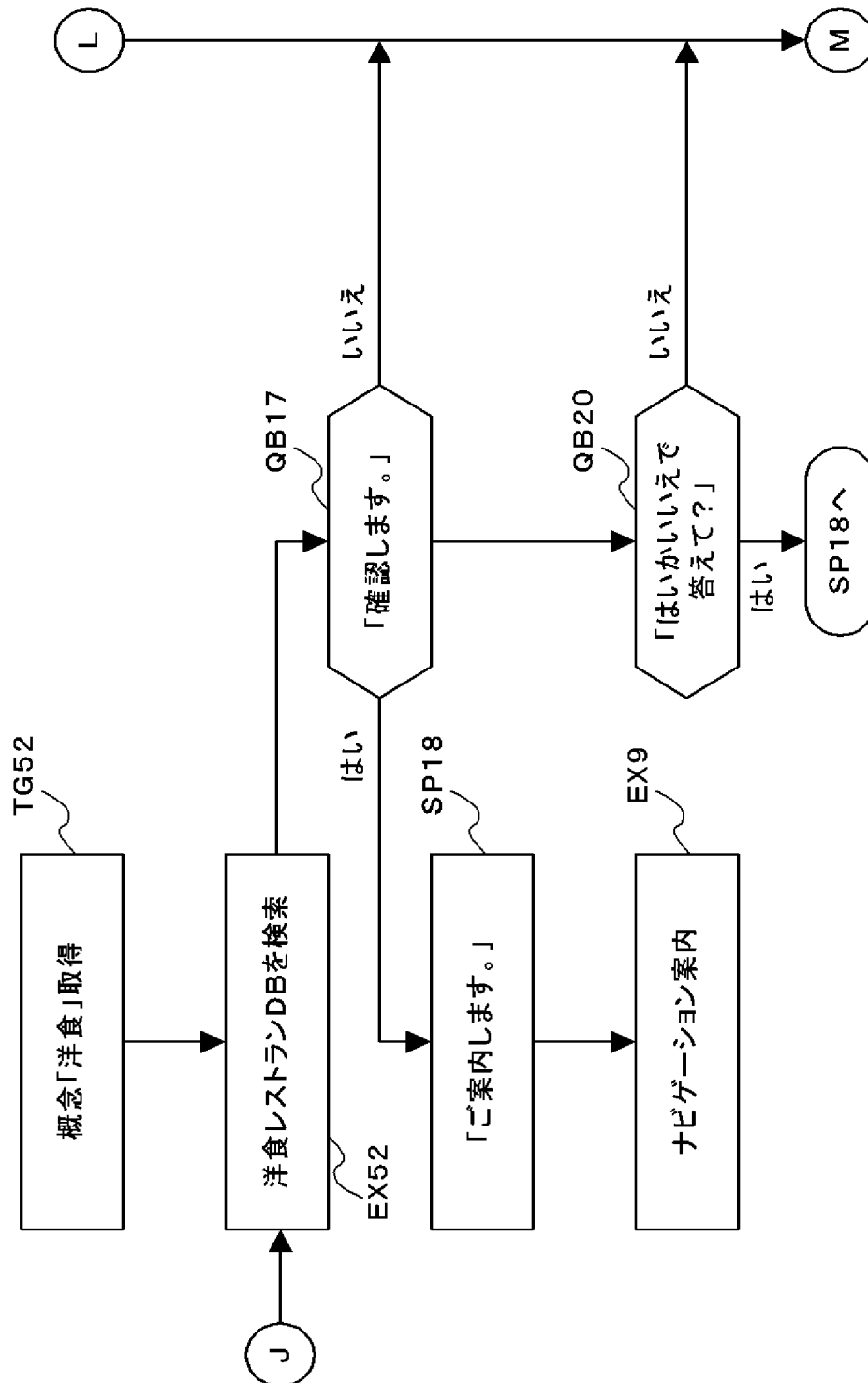
[図14]



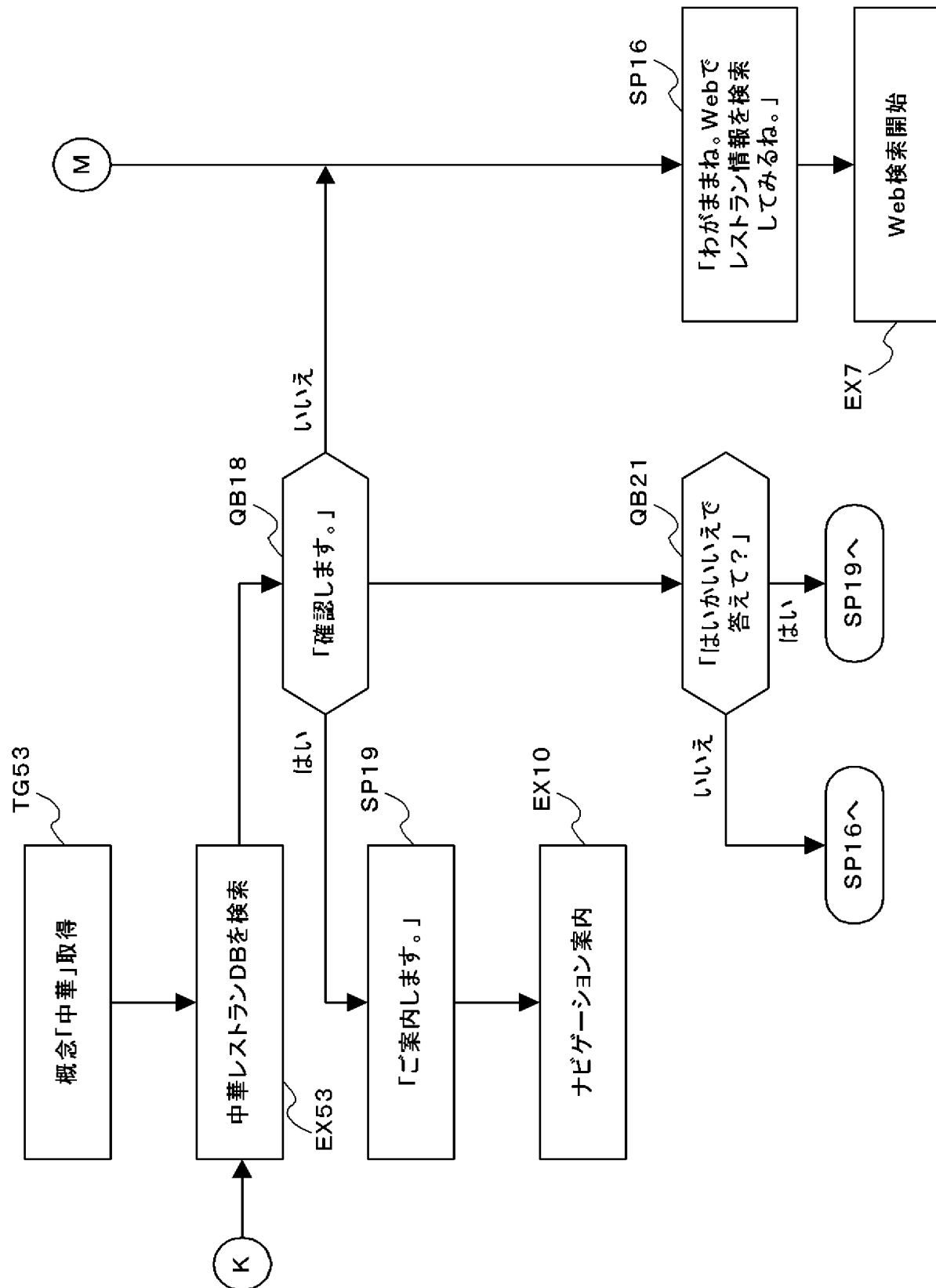
[図15]



[図16]



【図17】



[図18]

	「お腹」	「どこか」	「目的地」
「減った」	食事	食事	
「痛い」	病院	病院	
「設定」			ナビ設定
「行く」	食事	旅行	



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/019426

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.<sup>7</sup> G10L15/22, B60R16/02, G01C21/00, G08G1/0969, G09B29/10,  
G10L15/00, 15/18, 15/28

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.<sup>7</sup> G10L15/00-28

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2005  
Kokai Jitsuyo Shinan Koho 1971-2005 Toroku Jitsuyo Shinan Koho 1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2000-20086 A (Denso Corp.), 21 January, 2000 (21.01.00), Full text; Figs. 1 to 9 (Family: none)	1, 4-6, 11, 13-15, 20, 23-25, 31, 33-35, 40, 42-44, 49, 51-53 1-81
Y		
Y	JP 2001-249686 A (Matsushita Electric Industrial Co., Ltd.), 14 September, 2001 (14.09.01), Full text; Figs. 1 to 7 (Family: none)	1-81

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance  
"E" earlier application or patent but published on or after the international filing date  
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
"O" document referring to an oral disclosure, use, exhibition or other means  
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  
"&" document member of the same patent family

Date of the actual completion of the international search  
12 April, 2005 (12.04.05)

Date of mailing of the international search report  
26 April, 2005 (26.04.05)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/019426

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2002-341892 A (Matsushita Electric Industrial Co., Ltd.), 29 November, 2002 (29.11.02), Full text; Figs. 1 to 5 (Family: none)	1-81
Y	JP 2000-330588 A (Toshiba Corp.), 30 November, 2000 (30.11.00), Full text (particularly, Par. No. [0012]), Figs. 1 to 13 (Family: none)	2, 12, 21, 32, 41, 50
Y	JP 2003-44088 A (Sony Corp.), 14 February, 2003 (14.02.03), Full text (particularly, Par. No. [0003]), Figs. 1 to 7 & WO 2003/10756 A1 & US 2004/15344 A1	9, 18, 28, 38, 47, 56
Y	JP 8-339288 A (Canon Inc.), 24 December, 1996 (24.12.96), Full text; Figs. 1 to 5 (Family: none)	29

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. <sup>7</sup> G10L15/22, B60R16/02, G01C21/00, G08G1/0969, G09B29/10, G10L15/00, 15/18, 15/28		
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int.Cl. <sup>7</sup> G10L15/00-28		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2005年 日本国実用新案登録公報 1996-2005年 日本国登録実用新案公報 1994-2005年		
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP 2000-20086 A (株式会社デンソー) 2000.01.21 全文、図 1-9 (ファミリーなし)	1,4-6,11,13-15, 20,23-25,31,33- 35,40,42-44,49, 51-53
Y		1-81
Y	JP 2001-249686 A (松下電器産業株式会社) 2001.09.14 全文、図 1-7 (ファミリーなし)	1-81
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* / 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献		
国際調査を完了した日 12.04.2005	国際調査報告の発送日 26.04.2005	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号 100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 石丸 昌平 電話番号 03-3581-1101 内線 3539	5C 9559

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2002-341892 A (松下電器産業株式会社) 2002.11.29 全文、図 1-5 (ファミリーなし)	1-81
Y	JP 2000-330588 A (株式会社東芝) 2000.11.30 全文(特に【0012】)、図 1-13 (ファミリーなし)	2,12,21,32,41,50
Y	JP 2003-44088 A (ソニー株式会社) 2003.02.14 全文(特に段落【0003】)、図 1-7 & WO2003/10756A1 & US2004/15344A1	9,18,28,38,47,56
Y	JP 8-339288 A (キャノン株式会社) 1996.12.24 全文、図 1-5 (ファミリーなし)	29